1998

Visualizing silicon chemistry

Brett Myron Bode
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Visualizing silicon chemistry

by

Brett Myron Bode

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Physical Chemistry

Major Professor: Mark S. Gordon

Iowa State University

Ames, Iowa

1998

Graduate College

Iowa State University

This is to certify that the Doctoral dissertation of

Brett Myron Bode

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

There are many people I need to thank for their assistance and support in achieving this milestone in my career. First and foremost is my wife, Kim. Her love and support over that past several years have been invaluable to me as I worked towards this degree. She has also provided a helpful alternative viewpoint in many problems. In addition I would like to thank my son, Albert. who brightens every day, even if he does reduce the amount of sleep I receive at night. I would also like to thank my parents whose love and support have enabled me to reach this stage in my career. Their encouragement throughout my academic career has enabled me to choose the profession that provides me with the most challenge and enjoyment.

Of course one of the most important people supporting this research has been my advisor, Dr. Mark Gordon. Mark's advice and support have been invaluable to me as I have worked on my many research projects. I would also like to thank Dr. Mike Schmidt for many helpful discussions, particularly those involving Effective Core Potential theory and implementation.

ABSTRACT

The research presented in this thesis may be divided into three areas: transition metal catalysis, graphical user interfaces, and the derivation and application of effective core potentials.

Transition metal catalysis. The titanium catalyzed hydrosilation reaction has been examined in detail to determine a possible minimum energy reaction path. Two reactions and three catalysts were considered. In addition to a model system consisting of the reaction of silane with ethylene, the simplest known experimental reaction involving trichlorosilane and ethylene was also examined. In addition to a model catalyst of TiH_2, the expermental catalysts $TiCl_2$, and $Ti(C_5H_5)_2$ were considered. In all reactions studied the catalyst had a dramatic effect on the reaction system changing the overall reaction barrier from over 50 kcal/mol without the catalyst to a barrierless process with the catalyst. In addition to the overall reaction energies, several low energy intermediate structures were predicted which might be experimentally observable.

Graphical user interfaces. A description is given of a new graphical user interface for the GAMESS program. This interface includes many features useful for interpreting complex wavefunctions and reaction systems. These include the ability to animate reaction paths and normal modes of vibration, as well as the ability to view molecular orbitals, total electron densities, molecular electrostatic potentials, and density differences.

Effective core potentials. The derivation of an enhanced method for the computation of integrals involving effective core potentials is presented and has been implemented in the electronic structure code GAMESS. This method has helped produce a large reduction in the computational cost of ECPs. It has also enabled the implementation of analytical second

derivatives. The new method is also applied to the determination of the minimum energy structures of Si_8C_{12}, Ge_8C_{12} and Sn_8C_{12} which are main group analogs of the Ti_8C_{12} compounds (known as metcars). Relative energies, geometries, and vibrational frequencies are reported for several novel structures.

CHAPTER 1: GENERAL INTRODUCTION

General Overview

The hydrosilation reaction[1] is a general and industrially important method of producing silicon-carbon bonds and is important in the production of silicon-carbide precursors. As with many industrial reactions, very little was known about the mechanism for the reaction other than the basic parameters of temperature, reactants and reaction efficiency. The goal of the research on this reaction was therefore to answer the questions: "What is the minimum energy reaction path?" and "How does the catalyst lower the reaction energy barrier?"

Since the typical experimental system[1] is quite large, the initial series of calculations were performed on a model prototype system, designed by replacing the substituents with hydrogen atoms. The smaller model system enabled the assessment of several levels of theory such that we could choose the level of theory which gave reasonable accuracy, without being prohibitively expensive. Chapter 2 presents the results of this work. Once the level of theory necessary to correctly describe the reaction system was established, hydrosilation reaction systems that more closely model the experimental reaction were examined. These included the experimental reactants trichlorosilane and ethene,[2] as well as the experimental catalysts dichlorotitanium and dicyclopentadienyltitanium.[3] Chapter 3 presents the results of these calculations.

During the work on the hydrosilation reaction it became clear that the existing tools were inadequate for visualizing the results of complex reactions involving more than a few atoms. The tools that did exist were not designed for real time use and were limited to viewing 2D slices of orbitals and densities. This motivated the development of a program

that would run on any desktop Macintosh computer. The goal was to present the results of calculations in real time to the user without requiring any intermediate data processing steps between running the computation and viewing the results. In addition, it must be capable of visualizing very complex phenomena, including the animation of normal modes of vibration and reaction paths, and the computation and visualization of the molecular orbitals, total electron density and other complex properties of the wavefunction. The resulting program, MacMolPlt, is described in detail in Chapter 4.

Another important motivation in the development of MacMolPlt was its use as an educational tool. When students are beginning to learn the theory and the techniques of quantum chemistry, it can be difficult to also learn the intricacies of a large computational program. MacMolPlt helps reduce this hurdle by providing a simple interface to generate input files, the results of which can then be easily visualized. Having the ability to visualize orbitals, even basic atomic orbitals, with 3D models and real-time rotation can greatly enhance the understanding of orbital shapes and how orbitals interact to form bonds. Thus, MacMolPlt has proven to be a useful tool for a wide range of users from general chemistry students on up to experienced quantum chemists.

The hydrosilation reaction work also demonstrated the cost of accurate calculations. Since the most common computational techniques in quantum chemistry scale on the order of N^3 to N^7, where N is the size of the basis set, the cost of a calculation quickly becomes prohibitive as the size of the system increases. Thus, quantum chemists are constantly searching for ways to reduce the cost of a calculation without reducing the accuracy. One technique, known as the Effective Core Potential (ECP) method, replaces the core electrons with a potential. Thus, the ECP method directly reduces N and can dramatically speed up

calculations, particularly calculations involving fifth and sixth row elements. Unfortunately, the existing ECP implementation was inefficient, and lacked analytical second derivatives. To correct for these problems we developed and implemented a new ECP code including analytical first and second derivatives into the General Atomic and Molecular Electronic Structure System[4] (GAMESS) package.

Chapter 5 describes the new ECP implementation along with the determination of the minimum energy structures and energetics for 3 metallocarbohedrenes. Compounds with the formula M_8C_{12} have recently been reported for group 4 transition metals[5] (M = Ti, Zr, Hf). These compounds are proposed to have a cage structure similar to that found in fullerenes, except with a smaller size due the higher percentage of non-carbon atoms. It is reasonable to suspect that similar compounds exist using silicon, germanium, and tin instead of the transition metal, given the analogous electronic configuration, s^2p^2 versus s^2d^2. Thus, we have used ECPs to examine the structure and electronic configuration of Si_8C_{12}, Ge_8C_{12}, and Sn_8C_{12}. In fact, this work is an excellent example of how ECPs can reduce the cost of calculations involving heavy elements. Since all three compounds have exactly the same computational cost despite the fact that Ge and Sn have many more electrons than Si.

Dissertation Organization

This dissertation is made up of this general introduction, a chapter of general conclusions, and four papers that are either published, submitted for publication, or are being prepared for submission to peer reviewed journals with myself as the primary author.

Chapter 2 presents the results of calculations on a model system for the titanium catalyzed hydrosilation reaction. The results establish a base level of theory necessary for accurate results as well as presenting a model for the reaction path.

Chapter 3 expands on the results of Chapter 2 by considering the effects of the experimental substituents on the reactants and the catalyst. The results include a good model for the simplest experimentally observed hydrosilation reaction.

Chapter 4 describes the capabilities of the MacMolPlt program. This program provides a simple, easy to use interface to help interpret the results of complicated quantum mechanical calculations.

Chapter 5 presents a description and implementation of an improved method for computing integrals involving effective core potentials. Also included in Chapter 5 are the results of calculation predicting several minimum energy isomers for Si_8C_{12}, Ge_8C_{12} and Sn_8C_{12}.

Theory

The goal of quantum chemistry is to solve the complete Schrödinger equation[6] as a function of position, x, and time, t:

$$-\frac{\hbar}{i}\frac{\partial \Psi}{\partial t} = \hat{H}\Psi(x,t) \qquad (1)$$

where $\hat{H}$ is the Hamiltonian operator and Ψ is the total wavefunction. If this equation could be solved exactly for a system of interest, we would theoretically know all properties of that system as a function of time. In practice, since the potential energy is often time-independent, we can usually use the time-independent Schrödinger equation[7]:

$$\hat{H}\Psi = E\Psi \qquad (2)$$

Unfortunately this equation can be solved analytically only for one-electron atoms, such as the hydrogen atom. Thus in order to obtain a solution we must make a series of approximations. The first approximation is to assume that the effects of relativity are negligible. While this approximation is accurate for the lighter elements (such as the first 4 rows of the periodic table), it introduces significant errors into calculations involving heavier elements.

The next approximation made is the Born-Openheimer approximation.[8] This approximation assumes that the electrons are moving much faster than the nuclei. Therefore, the nuclear repulsion energy can be treated as a constant at a fixed geometry. Equation (1) can then be solved for the electronic Hamiltonian:

$$\hat{H}_{elec} = -\tfrac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{i}^{N}\sum_{A}^{N}\frac{Z_A}{r_{iA}} + \sum_{i>j}^{N}\frac{1}{r_{ij}} \qquad (3)$$

The first term represents the electronic kinetic energy, the second term the potential energy due to electron-nuclear attraction, and the third the potential energy due to the electron-electron repulsion. Unfortunately, the third term remains a three-body term, which makes equation (2) analytically unsolvable.

In Hartree-Fock theory[9,10] the n-electron system is separated into n one-electron equations given by:

$$\hat{F}\psi_i = \varepsilon_i\psi_i \qquad i = 1, 2, \ldots, n \qquad (4)$$

where ε_i are the orbital energies and $\hat{F}$ is the Fock operator given by:

$$\hat{F}(i) = -\tfrac{1}{2}\nabla_i^2 - \sum_{A=1}^{N}\frac{Z_A}{r_{iA}} + \hat{V}^{HF}(i) \qquad (5)$$

The first term, with the Laplacian operator ∇^2, involves differentiation with respect to the electronic coordinates giving the electron kinetic energy. The second term, with the atomic number of nucleus A, Z_A, and the distance between the ith electron and nucleus A, r_{iA}, gives the electron-nucleus attraction. The third term, $\hat{V}^{HF}$, represents the average potential experienced by electron i due to the other electrons in the molecule. This potential depends on the orbitals which may be represented by a linear combination of functions, χ_μ, centered on the nuclei:

$$\psi_i = \sum_i C_{\mu i} \chi_\mu \tag{6}$$

The χ_μ , referred to as basis functions, are usually gaussian functions, or linear combinations of gaussian functions, whose exponents have been predetermined from atomic and simple molecular calculations. Since the χ_μ are centered at the atomic nuclei, they can be considered the atomic orbitals. Thus, equation (6) represents a linear combination of atomic orbitals, or the LCAO approximation. When the LCAO approximation is introduced into the Hartree-Fock equations (equation (4)), the exact Hartree-Fock solution is obtained only in the limit of a complete basis set. However, it is possible to make a quite good approximation to this limit with a carefully selected finite set of orbitals that can also be systematically improved.

Because the Hartree-Fock equations are dependent on the average potential, $\hat{V}^{HF}$, they are also dependent on the initial orbitals. Therefore, they must be solved iteratively by making a guess for the initial orbitals, solving for the new orbitals, and repeating until a stable, self-consistent, solution is found. When this method is applied to a closed shell system with equal alpha and beta spatial orbitals it is referred to as restricted Hartree-Fock (RHF). Open shell calculations, in which the paired electrons have the same spatial orbitals and the

unpaired electrons each have their own alpha-spin orbital, are referred to as restricted open-shell Hartree-Fock (ROHF).

Unfortunately, Hartree-Fock theory does not account for the correlation of electrons, and thus allows electrons to come too close to each other. To correct for this problem several methods have been developed. The most commonly used correlation method is many-body perturbation theory. In perturbation theory, the effects of electron correlation are added back into the Hartree-Fock wavefunction by means of a perturbative treatment, and is thus appropriate only if the correlation effects are small. A form used extensively in this dissertation is Møller-Plesset second order perturbation theory[11] (MP2). Higher order perturbative methods are also available such as fourth order Møller-Plesset perturbation theory[12] (MP4), and coupled cluster theory[13] (CCSD(T)).

For many systems a single electron configuration is inadequate to fully describe the properties of the system. For such systems the multi-configuration self-consistent field (MCSCF) method[14] is needed. MCSCF variationally includes electron correlation and can be combined with a perturbative treatment such as multiconfigurational quasidegenerate second order perturbation theory[15] (MCQDPT) to provide accurate results.

Thus, given enough computer resources, we have the computational methods to solve most problems. The great problem is that in order to obtain accurate answers the basis set (equation (6)) must have many functions. Unfortunately, all of the computational methods mentioned have worse than linear scaling with respect to the number of basis functions. The Hartree-Fock method scales the best, but it still scales as approximately M^3, where M is the number of basis functions. Thus if M doubles the amount of work goes up by a factor of 8.

The methods for including correlation usually scale much worse, about M^5 for MP2 and M^7 for CCSD(T).

Given the poor scaling of these methods it would be advantageous if the basis set size could be reduced without reducing the accuracy of the calculation. For normal chemical processes (no core ionization), it is reasonable to expect that the atomic core orbitals will have constant energies and shapes from one system to another. Thus, one way to reduce the basis set size would be to replace the core orbitals with a potential. The potential is then designed to reproduce the effects of the core electrons on the valence orbitals. In other words it is an effective core potential (ECP).

The ECP can be derived by applying projection operators:

$$\hat{P} = \sum_{C=1}^{N} |\varphi_c\rangle\langle\varphi_c| \tag{7}$$

and the othogonal complement, $\left(1-\hat{P}\right)$, to the full set of molecular orbitals. By defining $\hat{P}$ to include all of the core orbitals we wish to remove, we ensure that when $\left(1-\hat{P}\right)$ is applied to the remaining valence orbitals the resulting orbitals are orthogonal to the core orbitals[16]. This allows us to rewrite the electronic hamiltonian (equation (3)) as[17]:

$$\hat{H}_{elec} = -\tfrac{1}{2}\sum_{v}^{N}\nabla_v^2 - \sum_{v}^{N}\sum_{A}^{N}\frac{Z_A}{r_{vA}} + U^{core}(v) + \sum_{v>j}^{N}\frac{1}{r_{ij}} \tag{8}$$

where the core orbitals have been moved into the U^{core} term and the remaining terms sum only over the valence orbitals. By inverting the modified Fock equation (equation (4)), the exact form of the core potential can be obtained.[18] The potential can then be fit using a linear combination of gaussians of the form:

$$r^{n-2}e^{-\alpha r^2} \tag{9}$$

where n represents the angular momentum of the core electrons removed by the potential. Similar to regular all electron basis sets, the coefficients and exponents of this linear combination are normally determined from calculations on individual atoms. However, since the ECP core is frozen, the fit can also be to a fully relativistic calculation on the atom. The resulting valence basis set and core potential can then correctly reproduce the contraction of the valence orbitals in heavy elements due to relativity. Thus, ECP calculations involving heavy elements can actually be more accurate than their all-electron counterparts.

Since U^{core} is an independent term in equation (8), we can easily separate out the ECP contribution to the total energy, which may be written as:

$$E_{ECP} = \langle \varphi_A | \hat{U}_{L+1} | \varphi_B \rangle + \sum_{lm}^{L} \langle \varphi_A | \hat{U}_{l-(L+1)} | lm \rangle \langle lm | \varphi_B \rangle \tag{10}$$

where L represents the maximum angular momentum function removed from the core. This results in the following types of angular and radial components[19]:

$$\Omega_{\lambda lm}^{abc} = \sum_{\mu=-\lambda}^{\lambda} Y_{\lambda\mu}(\Omega_k) \int d\Omega \hat{x}^a \hat{y}^b \hat{z}^c Y_{\lambda\mu}(\Omega) Y_{lm}(\Omega) \tag{11}$$

$$Q_{\lambda\lambda'}^{N}(k_A, k_B, \alpha) = \int_0^{\infty} dr r^N e^{-\alpha r^2} M_{\lambda}(k_A r) M_{\lambda'}(k_B r) \tag{12}$$

The first equation involving an integral over spherical harmonics, Y_{lm}, has a simple analytic solution. Unfortunately, the second equation involving an integral over two modified spherical bessel functions of the first kind, M_{λ}, is not as easy to solve. Chapter 5 will detail an improved solution to this integral.

In addition to the integrals required to compute the ECP contribution to the energy, it is also important to be able to compute analytic derivatives of those integrals. This is needed

in order to allow for efficient geometry optimizations and frequency calculations. However, since the form of the core potential is complex, the derivative of the core potential would be quite cumbersome to compute directly. Instead we make use of the principle of translational invariance[20] given by:

$$\left\langle \varphi_A \left| \frac{\partial \hat{U}_C}{\partial R_C} \right| \varphi_B \right\rangle = -\left[\left\langle \frac{\partial \varphi_A}{\partial R_A} \left| \hat{U}_C \right| \varphi_B \right\rangle + \left\langle \varphi_A \left| \hat{U}_C \right| \frac{\partial \varphi_B}{\partial R_B} \right\rangle \right] \tag{13}$$

to change derivatives of the potential into derivatives of the basis functions. Note that not only is the function being differentiated changed, but also the differentiated coordinate. For example, the ECP contribution to the gradient may be written as:

$$E^a_{ECP} = 2 \sum_{\mu} \sum_{\nu} \left\langle \varphi^{\alpha}_{\mu} \left| \delta_{AI} \sum_{C} \hat{U}_C - \hat{U}_A \right| \varphi_{\nu} \right\rangle \tag{14}$$

Since the derivative of a gaussian function is just another gaussian with a change in the angular momentum, the resulting derivatives are merely the sum of standard ECP integrals. For example, the derivative of a d-type gaussian results in the following integrals over s-type and f-type gaussians:

$$\langle d' | \hat{U} | \varphi \rangle = \langle s | \hat{U} | \varphi \rangle - 2\alpha \langle f | \hat{U} | \varphi \rangle \tag{15}$$

Since the computation of an ECP derivative requires the calculation of several integrals, it is important that the integral code be as fast as possible. Thus, we have also improved the speed of the integral code by a significant amount.

References

1. Barton, T. J.; Boudjouck, P. Organosilicon chemistry-a brief overview. In *Advances in Silicon-Based Polymer Science*; Advances in Chemistry Series No. 224; Ziegler, J.; Fearon, F. W., Eds.; American Chemical Society: Washington, DC, 1990; pp 3-46.
2. Speier, J. L. *Adv. Organomet. Chem.* **1979**, 17, 407.
3. Harrod, J. F.; Ziegler, T.; Tschinke, V. *Organometallics*, **1990**, 9, 897.
4. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery Jr., J. A.; The general atomic and molecular electronic structure system. *J. Comp. Chem.* **1993**, 14, 1347.
5. Guo, B. C.; Kerns, K. P.; Castleman, A. W., Jr. *Science* **1992**, 255, 1411.
6. Schrödinger, E. *Ann. Physik* **1926**, 81, 109.
7. (a) Schrödinger, E. *Ann. Physik* **1926**, 79, 734. (b) Schrödinger, E. *Ann. Physik* **1926**, 80, 437.
8. Born, M.; Openheimer, J. R. *Ann. Physik.* **1927**, 84, 457.
9. (a) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, 24, 89. (b) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, 24, 111. (c) Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, 24, 426. (d) Fock, V. *Physik* **1930**, 61, 126.
10. Szabo, A.; Ostlund, N. S. in *Modern Quantum Chemistry*, McGraw-Hill, New York, 1989.
11. (a) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, 46, 618. (b) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem.* **1976**, S10, 1.

12. Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, 72, 4244.

13. (a) Cizek, J. *J. Chem. Phys.* **1966**, 45, 4256. (b) Cizek, J. *Adv. Chem. Phys.* **1969**, 14, 35.

14. (a) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, 71, 41. (b) Ruedenburg, K.; Schmidt, M. W., Gilbert, M. M. *Chem. Phys.* **1982**, 71, 51. (c) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, 71, 65.

15. (a) Nakano, H. *J. Chem. Phys.* **1993**, 99, 7983. (b) Nakano, H. *Chem. Phys. Lett.* **1993**, 207, 372.

16. Weeks, J. D.; Rice, S. A. *J. Chem. Phys.* **1968**, 49, 2741.

17. Kahn, L.R.; Baybutt, P.; Trular, D. G. *J. Chem. Phys.* **1976**, 65, 3826.

18. (a) Kahn, L. R.; Goddard, W. A. *J. Chem. Phys.* **1972**, 56, 2685. (b) Melius, C. F.; Goddard, W. A. *Phys. Rev. A* **1974**, 10, 1528.

19. McMurchie, L.E.; Davidson, E. R. *J. Comp. Phys.* **1981**, 44, 289.

20. Kahn, L. R. *J. Chem. Phys.* **1981**, 75, 3962.

CHAPTER 2: THE CATALYZED HYDROSILATION REACTION

A paper published in and reprinted with permission from

Journal of the American Chemical Society **1998**, *120*, 1552-1555

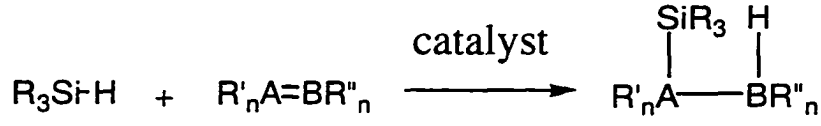
Brett M. Bode, Paul N. Day, Mark S. Gordon

Abstract

Ab initio electronic structure calculations using RHF, MP2, and CCSD(T) wavefunctions have been used to investigate a reaction path for the hydrosilation reaction catalyzed by divalent titanium (modeled by TiH_2). Optimized structures and energies are presented. All levels of theory predict a barrierless reaction path compared to a barrier of 78 kcal/mol for the uncatalyzed reaction. The use of correlated wavefunctions (MP2 or CCSD(T)) is required to obtain accurate structures and energies.

I. Introduction

The hydrosilation reaction is a general method of adding an Si-H bond across a C-C double bond. This method encompasses a wide variety of substituted alkenes, dienes, and alkynes leading to many different organosilicon products. Thus the method is very useful; indeed it is the second most important method of producing organosilanes on a large scale.[1] The general hydrosilation reaction may be written as:



One of the simplest examples known experimentally is the addition of trichlorosilane to ethylene, which will occur rapidly at room temperature and give nearly 100% yields with a variety of homogeneous transition metal based catalysts.[2]

Several analogous uncatalyzed reactions ($HSiCl_3$, SiH_4 + ethylene, SiH_4 + propene) were studied previously[3]; all were found to have large (≥54 kcal/mol) barriers. Thus, the catalyst is crucial in making the process economically viable. Industrially one active catalyst is believed to be a divalent Cp_2Ti species ($Cp = C_5H_5$).[1] The role of Cp_2Ti in catalyzing the polymerization of primary organosilanes has been studied by Harrod *et. al*.[4] These calculations suggest that Cp_2Ti exhibits a strong catalytic effect, but due to the size of the catalyst only selected points along the reaction path were studied in detail using double ζ and triple ζ quality basis sets and density functional theory. The nature of the bridging interactions between Ti and C in $Ti(C(Si(CH_3)_3)=C(C_6H_5)(CH_3))(Cp)^+_2$, an intermediate in a Ziegler catalyst system, has also been examined. The computed structure matched the experimental structure very well, even though the calculations employed fairly low levels of theory (RHF wavefunctions with a 3-21G basis set).[5]

There have been several recent studies on similar catalysts. A study of a silylene-bridged Cp_2Ti and its role in ethylene polymerization[6] employed RHF, MP2, and QCISD calculations with small basis sets (effective core potentials on the metal and 3-21G on the carbons and hydrogens). The results give a qualitative picture of the reaction path and several important structures along it, but the entire path was not examined. The Ziegler-Natta olefin polymerization process has been studied by several groups interested in the role of the $TiCl_2$ catalyst.[7,8,9] While these studies do provide a qualitative picture of the process, they all use relatively small basis sets and modest levels of theory (RHF and MP2).

This paper will consider the simplest prototypical example of a catalyzed hydrosilation reaction, in which A and B are carbon; R, R' and R" are hydrogen and the catalyst is TiH_2. The choice of reactants and catalyst allows mapping the entire reaction path at a high level of theory. Particularly, the choice of TiH_2 as the catalyst allows the use of high-level all-electron *ab inito* wavefunctions which would not be possible if the more complex catalysts such as $TiCl_2$ or $TiCp_2$ were used. Clearly, subsequent calculations will need to address the steric and electronic effects introduced by the Cl and Cp substituents. Nonetheless, TiH_2 provides an important baseline for comparison.

II. Computational Methods

The minimum energy reaction path connecting reactants to products was determined using all electron *ab initio* wavefunctions. The basis set used was a triple-ζ quality valance[10] plus 1 d-type polarization function on C, Si and 1 p-type polarization function on the hydrogens.[11]

The reaction path was determined by first optimizing the structures of the minima and transition states (TS) using analytic gradients and a modified Newton-Raphson algorithm. Each stationary point was confirmed by computing the matrix of energy 2nd derivatives, or hessian, to obtain the harmonic normal modes and corresponding frequencies (each minimum has zero and each transition state has one imaginary mode). The calculated frequencies were also used to obtain the harmonic zero-point energies used to convert energy differences to 0 K enthalpy differences. Finally the path connecting each TS to the nearest minima on each side of the TS was computed using the Gonzalez-Schlegel 2nd order intrinsic reaction coordinate path (IRC) method[12] with a step size of 0.3 $amu^{1/2}$-bohr for the first TS and 0.05

amu$^{1/2}$-bohr for the second TS.

The path was initially optimized at the restricted Hartree-Fock (RHF) level of theory, then refined using Møller-Plesset second order perturbation theory (MP2). Single-point energies were computed at the MP2 optimized stationary points using coupled cluster singles and doubles plus perturbative triples (CCSD(T)).

The GAMESS[13] program was used for all of the RHF calculations and a portion of the MP2 optimizations. The Gaussian 92 suite of programs[14] was used for the remainder of the MP2 calculations and the CCSD(T) calculations.

III. Results and Discussion

Figure 1 shows the energy profile of the proposed catalyzed reaction. The zero of energy on the curve for each level of theory is the sum of the reactant energies at that level of theory (structures **a**, **b**, and **c** in Fig. 2). The MP2 structures at each stationary point are given in Fig. 2[15]. The MP2 and CCSD(T) total energies and the MP2 vibrational zero point energy (ZPE) corrections for each geometry point marked in Fig. 1 and 2 are available as an appendix to this chapter. MP2 and CCSD(T) ZPE corrected energies are listed relative to the zero of energy in Table 1.

It is important to note that all points on the energy plot in Fig. 1 lie below the energy of the reactants, in contrast to the large barrier in the uncatalyzed reaction. Note also that there are large differences between the SCF and MP2 energy profiles, while the differences between MP2 and CCSD(T) are much smaller. So, electron correlation is essential for a correct description of this reaction surface, and MP2 is qualitatively correct.

There are two possibilities for the first step of the reaction, both of which are

barrierless processes. The first, and more exothermic, is to add the TiH_2 catalyst across the ethylene double bond to form the three membered ring compound shown in Fig. 2d. This process is downhill in energy by 61.9 (53.4) kcal/mol at the ZPE corrected MP2 (CCSD(T)) level of theory. Note that, based on the large exothermicity and the large (0.016Å) increase in the CC bond length, structure **d** is a three-membered ring, not a π complex. Silane will then add to form the complex depicted in Fig. 2e. This second barrierless addition is downhill by 6.5 (6.0) kcal/mol.

The electronic structure of TiH_2 was considered in detail previously.[16] Like CH_2, the ground state is a triplet, and the lowest singlet state is 21 kcal/mol higher in energy. Since TiH_2 has an electronic structure similar to singlet CH_2 or SiH_2, (s^2d^2 vs. s^2p^2) a reasonable expectation is that an alternative mechanism would start with an insertion of TiH_2 into an Si-H bond of silane. Indeed, this occurs with no barrier to produce structure **d'**, a Ti-Si analog of ethane. This step is downhill by 31.9 (27.8) kcal/mol. When ethylene is added to this compound, it rearranges with no barrier to the same structure as in Fig. 2e.

So, whether the TiH_2 catalyst adds to ethylene or silane initially, the net result after the two reactants and the catalyst have been added together is the formation of compound **e**, with no intervening barrier. The overall exothermicity to this point is 68.4 (59.4) kcal.mol. This very large drop in energy drives the entire reaction path down in energy. In fact the reaction path is forced down enough that all subsequent points are below the reactants in energy.

Recall that the final desired product is ethylsilane. So, starting from compound 2e, the silyl group needs to migrate to the nearest (α) carbon, and a hydrogen needs to be transferred to the adjacent (β) carbon, with the ultimate removal of the TiH_2 catalyst. Therefore, the next

step in the reaction is to transfer a H from the complexed silane to the Ti and to attach the Si to the α C.

The first part of this step involves rotation of the silane such that there are 2 bridging hydrogens between the Si and Ti; that is e->g via f in Figs. 1 and 2. At the MP2 level there is a small barrier (at structure f) of 1.7 kcal/mol to this process, but after the ZPE corrections are added the barrier disappears. The reaction then proceeds through transition state h with a barrier of 5.3 (8.2) kcal/mol, leading to the four-membered ring shown in Fig. 2i. This ring is 5.9 (6.6) kcal/mol below the TS h. The four membered ring can be opened up by breaking the Si-Ti bond to give the compound **k**. The TS for this step is shown in Fig. 2j. It has a barrier height of 3.2 (1.2) kcal/mol. Compound **k** is 1.4 (2.1) kcal/mol below the TS j.

The final step in the process is to regenerate the catalyst by transfer of a hydrogen from Ti to C and elimination of TiH_2. The transition state for this process is shown in Fig. 2l; the associated barrier height is 31.9 (25.7) kcal/mol. This TS is still 33.5 (31.1) kcal/mol lower in energy than the initial reactants. The IRC from this TS leads to the structure shown in Fig. 2m which is 5.2 (5.9) kcal/mol below the TS. The structure shown in Fig. 2m is not a stationary point, but it illustrates that the reaction path goes through a structure in which the TiH_2 is complexed to two hydrogens. Optimization from this point leads to the insertion of TiH_2 into an Si-H bond, as shown in Fig. 2n. The insertion product is 25.7 (22.2) kcal/mol below the TS in energy. However, the TiH_2 in Fig. 2m is not tightly bound to the ethylsilane as evidenced by both the relatively long Ti-Si and Ti-H bond distances and the fact that it is a modest 11.0 (11.6) kcal/mol in energy below separated products. Thus we do not expect a transition state for the process of simply abstracting the TiH_2 to form separated products. This can occur readily due to the excess energy available to the system, since the separated

products are 28.0 (25.4) kcal/mol below the reactants in energy. Note that the *reverse* barriers for elimination of singlet CH_2 from methane and singlet SiH_2 from silane are ~ zero.[17] It is also much less likely that the catalyst will insert when both the catalyst and the silyl group have more bulky substituents such as Cl or Cp rings.

Once TiH_2 is removed, the process is complete with ethylsilane as the product. The overall process is exothermic by 28.0 (25.4) kcal/mol at the ZPE corrected MP2 (CCSD(T)) level of theory. This compares with the value of 29.1 kcal/mol computed by Day and Gordon[3] at the MP2/6-311G(d,p) level of theory and a value of 27.4 kcal/mol computed by McDouall *et. al.*[18] at the MP4/6-31G(d)//HF/3-21G level of theory. There does not seem to be a good experimental ΔH_0 for this reaction, but we can estimate the value to be 28.9 kcal/mol based on the experimental heats of formation for ethene and silane,[19] and the best previous theoretical heat of formation for ethylsilane.[20]

The driving force for the entire reaction comes in the first two steps with the formation of the compound shown in Fig. 2e which is 68.4 (59.4) kcal/mol below the reactants in energy and is the global minimum on the reaction surface. The reasons this structure is so stable are illustrated by the first two steps in the reaction. In the first step the electron deficient TiH_2 adds to the ethene across the π bond in much the same manner as the addition of CH_2 to ethylene to form cyclopropane. The second step is much less exothermic and is driven mostly by the electrostatic attraction between the positively charged titanium (+0.83) and the negative hydrogen (-0.12) on the silicon.

IV. Conclusions

The results presented here clearly show that divalent titanium is an effective catalyst for the hydrosilation reaction. Further work, currently in progress, will examine the effect of substitution on catalyst and reactants, but even the simple model catalyst, TiH_2, used here clearly shows strong catalytic behavior. The overall catalyzed reaction has no net barrier, because of the very stable cyclic $TiH_2CH_2CH_2$ intermediate. However, the energy profile of the multistep process (Fig. 1) does offer the possibility of finding some of the intermediate structures if the process was carried out at low temperature.

Acknowledgements

The work described in this paper was supported by grants from Iowa State University in the form of a Department of Education GAANN fellowship awarded to BMB and in the form of a grant to purchase computers used in this project, the National Science Foundation (CHE-9633480) and the Air Force Office for Scientific Research (F49-620-95-1-0073). The computations were performed in part on computers provided by Iowa State University and through a grant of computer time at the San Diego Supercomputer Center.

References

1. Barton, T.J.; Boudjouck, P "Organosilicon chemistry - a brief overview", in *Advances in Silicon-Based Polymer Science*, Advances in Chemistry Series No. 224, Ziegler J.; Fearon, F. W. G. (eds) American Chemical Society, Washington, DC. **1990**, 3-46.
2. Speier, J. L. *Adv. Organomet. Chem.* **1979**, *17*, 407-447.

3. Day, P.N.; Gordon, M.S. *Theor. Chim. Acta.* **1995**, *91*, 83-90.

4. Harrod, J.F.; Ziegler, T.; Tschinke, V. *Organometallics* **1990**, *9*, 897-902.

5. Koga, N.; Morokuma, K. *J. Am. Chem. Soc.* **1988**, *110*, 108-112.

6. Yoshida, T.; Koga, N.; Morokuma, K. *Organometallics* **1995**, *14*, 746-758.

7. Jensen, V. R.; Børve, K. J.; Ystenes, M. *J. Am. Chem. Soc.* **1995**, *117*, 4109-4117.

8. Sakai, S. *J. Phys. Chem.* **1994**, *98*, 12053-12058.

9. Kawamura-Kuribayashi, H.; Koga, N.; Morokuma, K. *J. Am. Chem. Soc.* **1992**, *114*, 2359-2366.

10. For H, C: Dunning, T.H. *J. Chem. Phys.* **1971**, *55*, 716-723.; for Si: McLean, A.D.; Chandler, G.S. *J. Chem. Phys.* **1980**, *72*, 5639-5648.; for Ti: Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033-1036. This has been extended to triple-ζ quality by using d functions (developed by: Rappe, A.K.; Smedley, T.A.; Goddard, W.A., III *J. Phys. Chem.* **1981**, *85*, 2607-2611.), deleting Wachters' most diffuse s function $\zeta = 0.0333$, adding an s function with $\zeta = 0.209$ to better describe the 3s-4s region, and adding two p functions $\zeta = 0.1506$ and 0.0611 to allow for a 4p.

11. The exponents used were: H $\zeta_p = 1.0$; C $\zeta_d = 0.72$; Si $\zeta_d = 0.388$

12. (a) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523-5527. (b) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1991**, *95*, 5853-5860.

13. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. Jr. The general atomic and molecular electronic structure system. *J. Comp. Chem.* **1993**, *14*, 1347-1363.

14. Gaussian 92 rev C, Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill PM. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogie, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. L.; Defrees, D. J.; Baker, J.; Stewart JJ. P.; Pople, J. A. Gaussian Inc., Pittsburg PA, 1992.

15. The full set of cartesian coordinates for each structure is available as an appendix and on the WWW site: http://www.msg.ameslab.gov/

16. Kudo, T.; Gordon, M.S. *J. Chem. Phys.* **1995**, *102*, 6806-6811.

17. Gordon, M. S.; Gano, D. R.; Binkley, J. S.; Frisch, M. J. *J. Am. Chem. Soc.*, **1986**, *108*, 2191-2195.

18. McDouall, J. J. W.; Schlegel, H. B.; Francisco, J. S. *J. Am. Chem. Soc.* **1989**, *111*, 4622-4627.

19. Wagman, D. D.; Evans, W. H.; Parker, V. B.; Schumm, R. H.; Halow, I.; Bailey, S. M.; Churney, K. L.; Nuttall, R. L. *J. Phys. Chem. Ref. Data Suppl.* **1982**, *11*, 2.

20. Gordon, M. S.; Boatz, J. A.; Walsh, R. *J. Phys. Chem.* **1989**, *93*, 1584-1585.

Table 1: MP2 and CCSD(T) relative energies (kcal/mol) with ZPE correction

Geometry point	MP2 + MP2 ZPE	CCSD(T) + MP2 ZPE
a+b+c (reactants)	0	0
d'+b	-31.1	-27.8
d+c	-61.9	-53.4
e	-68.4	-59.4
f	-66.6	-57.8
g	-66.6	-57.4
h	-61.3	-49.2
i	-67.2	-55.8
j	-64.0	-54.6
k	-65.4	-56.8
l	-33.5	-31.1
m	-39.0	-37.0
n	-59.2	-53.3
o+a (products)	-28.0	-25.4

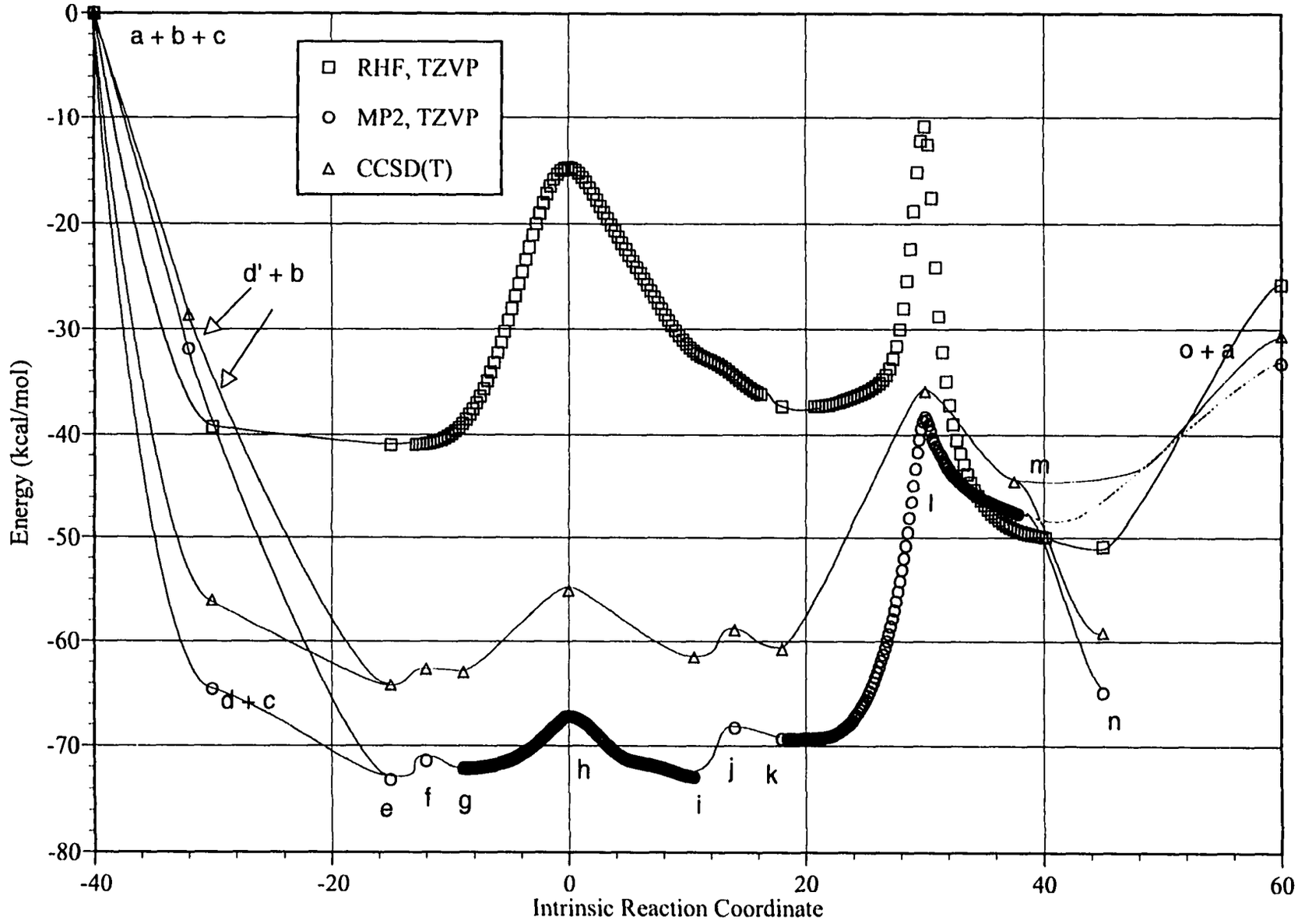


Figure 1: RHF, MP2, and CCSD(T) Energies

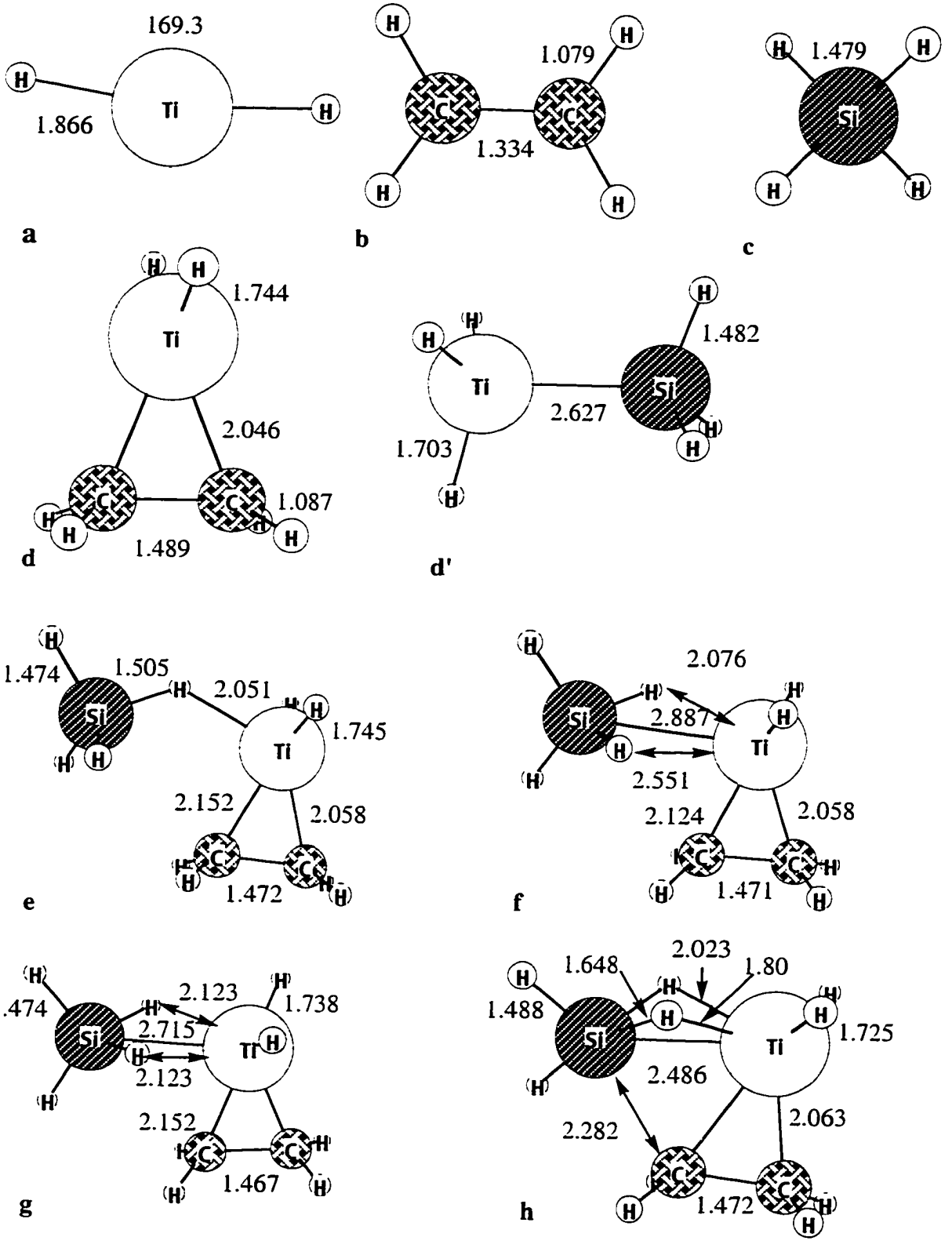
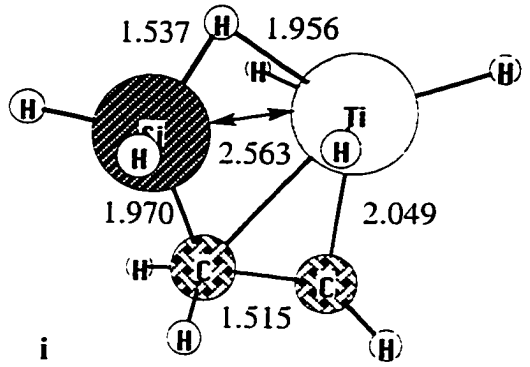
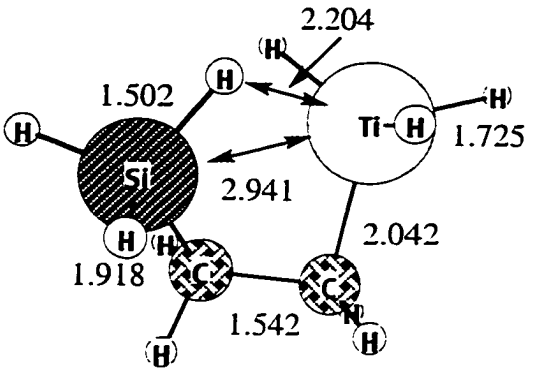


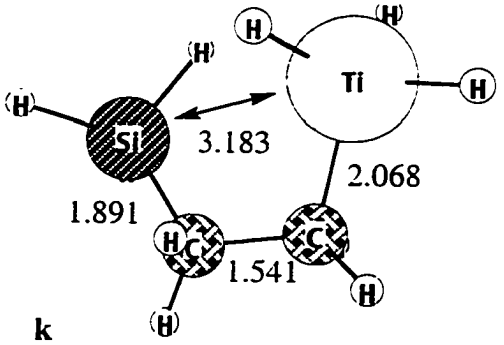
Figure 2: MP2 structures along the minimum energy reaction path



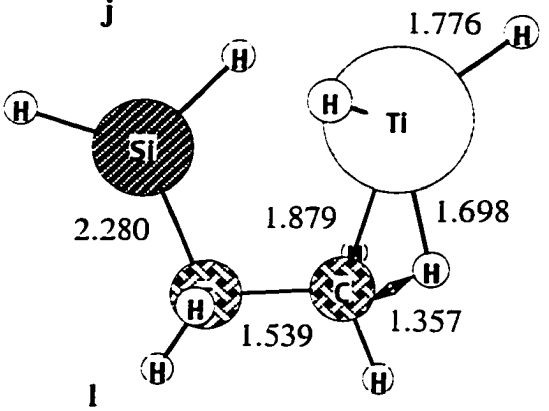
i



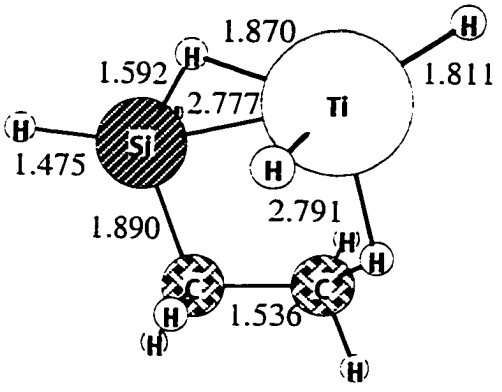
j



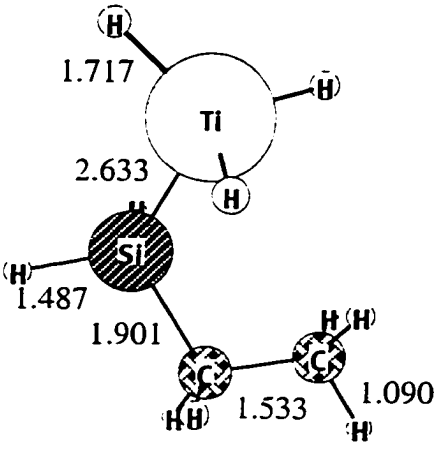
k



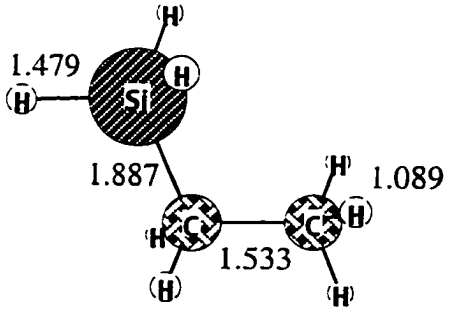
l



m



n



o

Figure 2: Continued.

Appendix

MP2 and CCSD(T) total energies (in Hartrees)

Geometry point	MP2	CCSD(T)	MP2 ZPE
a	-849.523689	-849.558595	0.007380
b	-78.350167	-78.388712	0.051503
c	-291.367241	-291.396783	0.032340
d'	-1140.941807	-1141.000986	0.041009
d	-927.976820	-928.036620	0.063141
e	-1219.357669	-1219.446391	0.098847
f	-1219.354932	-1219.443927	0.098939
g	-1219.356058	-1219.444445	0.100075
h	-1219.348157	-1219.431893	0.100614
i	-1219.357310	-1219.442124	0.100291
j	-1219.349955	-1219.437991	0.098052
k	-1219.351630	-1219.440864	0.097549
l	-1219.302084	-1219.401222	0.098800
m	-1219.314969	-1219.414914	0.103070
n	-1219.344683	-1219.438353	0.100531
o	-369.770333	-369.834273	0.092168

Fig. 2a:

Ti 22.0 0.00000 0.00000 0.01455

H 1.0 0.00000 1.85751 -0.16011

H 1.0 0.00000 -1.85751 -0.16011

Fig 2b:

C 6.0 0.00000 0.00000 0.00000

C 6.0 0.00000 0.00000 1.33420

H	1.0	0.92095	0.00000	-0.56413
H	1.0	-0.92095	0.00000	-0.56413
H	1.0	0.92095	0.00000	1.89833
H	1.0	-0.92095	0.00000	1.89833

Fig 2c:

Si	14.0	0.00000	0.00000	0.00000
H	1.0	0.85112	0.85112	0.85112
H	1.0	-0.85112	-0.85112	0.85112
H	1.0	-0.85112	0.85112	-0.85112
H	1.0	0.85112	-0.85112	-0.85112

Fig 2d:

Ti	22.0	-0.75328	-0.00028	0.00010
C	6.0	1.15258	0.74494	-0.00086
C	6.0	1.15389	-0.74442	0.00001
H	1.0	-1.57672	0.00186	1.53734
H	1.0	-1.58794	-0.00135	-1.53226
H	1.0	1.47334	1.24414	-0.91146
H	1.0	1.47338	1.24461	0.90945
H	1.0	1.47648	-1.24261	0.91041
H	1.0	1.47472	-1.24349	-0.91062

Fig 2e:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.47230

Ti	22.0	1.94598	0.00000	0.66971
H	1.0	-0.29841	-0.91292	-0.50901
H	1.0	-0.28919	0.91361	-0.51308
H	1.0	-0.30839	0.91870	1.96470
H	1.0	-0.35958	-0.89759	1.96888
H	1.0	2.73996	-1.53409	0.41951
H	1.0	2.72915	1.54487	0.45418
Si	14.0	2.12669	-0.32911	3.80391
H	1.0	1.20298	0.74146	4.19473
H	1.0	2.76619	-0.00219	2.48145
H	1.0	3.29348	-0.36718	4.70299
H	1.0	1.48851	-1.64853	3.73042

Fig 2f:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.46720
Ti	22.0	1.97067	0.00000	0.60347
Si	14.0	2.30005	-0.00351	3.29852
H	1.0	-0.27693	-0.91406	-0.51838
H	1.0	-0.27712	0.91390	-0.51856
H	1.0	-0.34887	-0.90523	1.95508
H	1.0	-0.34806	0.90568	1.95504
H	1.0	2.41857	-1.14752	2.33255
H	1.0	2.70599	-1.48744	0.08798

H	1.0	2.70536	1.48863	0.09053
H	1.0	3.53202	-0.00516	4.10688
H	1.0	1.11085	-0.00384	4.15427
H	1.0	2.42006	1.14291	2.33554

Fig 2g:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.47080
Ti	22.0	1.95546	0.00000	0.64183
Si	14.0	2.31328	-0.45634	3.47019
H	1.0	-0.28919	-0.91510	-0.51043
H	1.0	-0.28039	0.91144	-0.52092
H	1.0	-0.38705	-0.88643	1.96349
H	1.0	-0.29691	0.92467	1.95906
H	1.0	2.15538	-1.63643	2.58905
H	1.0	2.84038	-1.42768	0.16633
H	1.0	2.61907	1.56791	0.30430
H	1.0	3.53009	-0.59407	4.28719
H	1.0	1.13594	-0.15624	4.29001
H	1.0	2.62211	0.64869	2.49765

Fig 2h:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.46967
Ti	22.0	2.02424	0.00000	0.40507

Si	14.0	1.80765	-0.17259	2.87880
H	1.0	-0.28995	-0.91347	-0.51179
H	1.0	-0.30152	0.91703	-0.49820
H	1.0	-0.46390	-0.86691	1.93421
H	1.0	-0.33371	0.92523	1.92836
H	1.0	2.34942	-1.19408	1.72081
H	1.0	2.69724	-1.40954	-0.32716
H	1.0	2.73766	1.50968	-0.07396
H	1.0	2.67777	-0.90378	3.83621
H	1.0	0.85379	0.48616	3.81066
H	1.0	2.59709	0.90750	2.12531

Fig 2i:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.51570
Ti	22.0	2.04070	0.00000	0.18316
Si	14.0	1.54489	-0.42358	2.66192
H	1.0	-0.39984	-0.89015	-0.47502
H	1.0	-0.33956	0.92668	-0.45531
H	1.0	-0.62185	-0.82319	1.87698
H	1.0	-0.35549	0.94501	1.91810
H	1.0	1.93691	-1.56678	0.94415
H	1.0	2.87280	-0.24317	-1.34381
H	1.0	2.09440	1.62669	0.77810

H	1.0	1.56461	-1.76676	3.25667
H	1.0	1.42732	0.53740	3.78398
H	1.0	2.83123	0.05677	1.97176

Fig 2j:

C	6.0	0.46725	-1.51614	0.31695
C	6.0	-0.90428	-1.17315	-0.30196
Ti	22.0	1.13445	0.38133	-0.03025
Si	14.0	-1.67358	0.56202	0.04934
H	1.0	0.41127	-1.70885	1.38990
H	1.0	0.99250	-2.31658	-0.19858
H	1.0	-1.65722	-1.83623	0.13305
H	1.0	-0.90245	-1.34338	-1.37624
H	1.0	1.38451	1.15190	1.49626
H	1.0	2.68112	0.28213	-0.79289
H	1.0	0.32219	1.01616	-1.42968
H	1.0	-2.61542	0.42162	1.17883
H	1.0	-2.34323	1.16549	-1.10764
H	1.0	-0.58419	1.44949	0.60678

Fig 2k:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.54080
Ti	22.0	1.99890	0.00000	-0.53052
Si	14.0	1.01141	1.48621	2.12515

H	1.0	-0.47177	0.89608	-0.41784
H	1.0	-0.47653	-0.88238	-0.43622
H	1.0	-1.01677	0.06060	1.93805
H	1.0	0.42742	-0.92420	1.93226
H	1.0	2.43231	1.48604	-1.27898
H	1.0	2.10495	-1.15771	-1.79665
H	1.0	2.95873	-0.68521	0.73495
H	1.0	2.11026	1.62321	1.11023
H	1.0	1.62396	1.34860	3.45777
H	1.0	0.22215	2.72935	2.03312

Fig 2l:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.53940
Ti	22.0	2.26199	0.00000	-0.28365
Si	14.0	0.98399	1.52691	2.02113
H	1.0	-0.19491	0.99029	-0.43160
H	1.0	-0.81226	-0.63076	-0.36898
H	1.0	-1.00627	-0.01126	1.96467
H	1.0	0.49796	-0.89862	1.91565
H	1.0	2.90730	0.74599	-1.76068
H	1.0	0.83756	-0.84884	-0.64791
H	1.0	3.16153	-1.01623	0.86297
H	1.0	2.19694	1.50541	1.10849

H	1.0	1.51048	1.60041	3.39357
H	1.0	0.26208	2.75549	1.64571

Fig 2m:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.53692
Ti	22.0	2.69344	0.00000	0.08633
Si	14.0	1.10140	1.40473	2.14497
H	1.0	-0.04045	0.99948	-0.43191
H	1.0	-0.85125	-0.55266	-0.39632
H	1.0	-1.01490	0.11485	1.92151
H	1.0	0.36845	-0.95477	1.91449
H	1.0	3.31671	0.73100	-1.46679
H	1.0	0.83295	-0.57519	-0.46180
H	1.0	2.98004	-1.48783	1.10494
H	1.0	2.52619	1.27445	1.58528
H	1.0	1.32002	1.43052	3.60087
H	1.0	0.58887	2.69834	1.65661

Fig 2n:

C	6.0	2.22770	-1.31559	0.24591
C	6.0	2.10510	0.04018	-0.45865
Ti	22.0	-1.61814	-0.27378	-0.04059
Si	14.0	0.63320	1.07832	0.14873
H	1.0	2.29794	-1.19808	1.32688

H	1.0	3.11462	-1.85287	-0.08776
H	1.0	3.01469	0.62483	-0.30460
H	1.0	2.00714	-0.10444	-1.53570
H	1.0	-1.45196	-1.41704	1.22798
H	1.0	1.36534	-1.95041	0.04064
H	1.0	-1.46083	-0.92722	-1.62176
H	1.0	-2.84307	0.90975	0.17370
H	1.0	0.52275	2.33646	-0.63579
H	1.0	0.83760	1.45059	1.57492

Fig 2o:

C	6.0	0.00000	0.00000	0.00000
C	6.0	0.00000	0.00000	1.53350
Si	14.0	1.73746	0.00000	2.26971
H	1.0	1.68863	0.00050	3.74790
H	1.0	2.47762	-1.20015	1.81872
H	1.0	2.47545	1.20002	1.81942
H	1.0	-1.01460	-0.00027	-0.39616
H	1.0	0.50793	0.87970	-0.39389
H	1.0	0.50827	-0.87947	-0.39395
H	1.0	-0.53698	-0.87394	1.90667
H	1.0	-0.53742	0.87367	1.90667

CHAPTER 3: THE CATALYZED HYDROSILATION REACTION: SUBSTITUENT EFFECTS

A paper accepted for publication in *Theoretical Chemical Accounts*

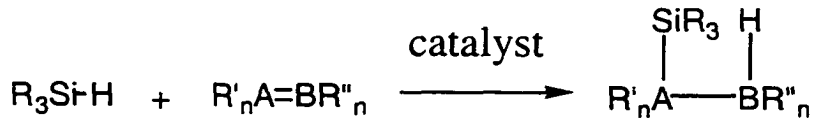
Brett M. Bode, Mark S. Gordon

Abstract

Ab initio electronic structure calculations using MP2 wavefunctions have been used to investigate a reaction path for the hydrosilation reaction catalyzed by divalent titanium (modeled by TiH_2, $TiCl_2$, and $Ti(C_5H_5)_2$). Optimized structures and energies are presented. All model reactions predict a barrierless reaction path compared to a barrier of 78 kcal/mol for the uncatalyzed reaction.

I. Introduction

The hydrosilation reaction is a general method for adding an Si-H bond across a C-C double bond. This method encompasses a wide variety of substituted alkenes, dienes, and alkynes leading to many different organosilicon products. Thus the method is very useful; indeed it is the second most important method of producing organosilanes on a large scale [1]. The general hydrosilation reaction may be written as:



One of the simplest examples known experimentally is the addition of trichlorosilane to ethylene, which will occur rapidly at room temperature and give nearly 100% yields with a variety of homogeneous transition metal based catalysts [2].

Several analogous uncatalyzed reactions ($HSiCl_3$, SiH_4 + ethylene, SiH_4 + propene) were studied previously [3]; all were found to have large (≥54 kcal/mol) barriers. Thus, the catalyst is crucial in making the process economically viable. Industrially one active catalyst is believed to be a divalent $TiCp_2$ species (Cp = C_5H_5). The role of $TiCp_2$ in catalyzing the polymerization of primary organosilanes has been studied by Harrod *et. al* [4]. These calculations suggest that $TiCp_2$ exhibits a strong catalytic effect, but due to the size of the catalyst only selected points along the reaction path were studied in detail using double ζ and triple ζ quality basis sets and density functional theory. The nature of the bridging interactions between Ti and C in $Ti(C(Si(CH_3)_3)=C(C_6H_5)(CH_3))(Cp)^+_2$, an intermediate in a Ziegler catalyst system, has also been examined. The computed structure matched the experimental structure very well, even though the calculations employed fairly low levels of theory (RHF wavefunctions with a 3-21G basis set) [5].

There have been several recent studies on similar catalysts. A study of a silylene-bridged $TiCp_2$ and its role in ethylene polymerization [6] employed RHF, MP2, and QCISD calculations with small basis sets (effective core potentials on the metal and 3-21G on the carbons and hydrogens). The results give a qualitative picture of the reaction path and several important structures along it, but the entire path was not examined. The Ziegler-Natta olefin polymerization process has been studied by several groups interested in the role of the $TiCl_2$

catalyst [7,8,9]. Similar structures were also found in a study of the McMurry reaction involving the $TiCl_2$ catalyzed reductive coupling of carbonyl compounds [10]. While these studies do provide a qualitative picture of the process, they all use relatively small basis sets and modest levels of theory (RHF and MP2).

In a previous paper [11] we considered the simplest prototypical example of a catalyzed hydrosilation reaction, in which A and B are carbon; R, R' and R" are hydrogen and the catalyst is TiH_2. The choice of reactants and catalyst allowed mapping the entire reaction path at a high level of theory (MP2 geometries plus CCSD(T) energetics). From this baseline work it was determined that the MP2 level of theory was adequate to determine both the reaction energies and the optimized geometries. Preliminary calculations on the effect of Cl-substitution on the hydrosilation reaction have recently been reported [12].

In this paper, we will examine the similarities and differences between the model system in our previous work and the actual experimental systems: the reaction studied will be ethylene + trichlorosilane, the simplest experimental system. The catalysts will include the model catalyst from our previous work, TiH_2, the catalyst used in model studies of many similar reactions, $TiCl_2$, and finally the catalyst believed to be involved in the experimental reaction system, $TiCp_2$.

II. Computational Methods

In our previous work the minimum energy reaction path connecting reactants to products was determined using all electron *ab initio* wavefunctions with a basis set of triple-ζ

plus polarization quality. To examine the effects of substituents we required a more computationally efficient basis set capable of similar accuracy to our previous work.

We first tried using an SBKJC Effective Core Potential [13] (ECP) basis on C, Si, Cl, Ti and a 6-31G(p) basis set on hydrogen. The SBKJC basis set was extended with one d-type polarization function on each C, Si, and Cl [14]. This basis set was evaluated by reoptimizing the stationary points in the reaction path from our previous work. As shown in Figure 1, the SBKJC basis set tracks the triple ζ basis, but it overestimates the exothermicity by 7-8 kcal/mol over most of the reaction path leading to an rms deviation of 6.7 kcal/mol. However, it was found that the rms difference drops to 0.5 kcal/mol when the C ECP basis set is replaced with a 6-31G(d) basis set (see Figure 1). Thus we have used a 6-31G(d,p) [15] basis set on H and C, and the SBKJC ECP basis on Si, Cl, and Ti extended with one d-type polarization function on Si and Cl.

All calculations were performed using closed shell Møller-Plesset second order perturbation theory (MP2) as implemented in the GAMESS [16] program. The reaction paths involving TiH_2 and $TiCl_2$ were determined by first optimizing the structures of the minima and transition states (TS) using analytic gradients and a modified Newton-Raphson algorithm. Each stationary point was confirmed by computing the matrix of energy second derivatives (hessian), to obtain the harmonic normal modes and corresponding frequencies (each minimum has zero and each transition state has one imaginary mode). The calculated frequencies also provide the harmonic zero-point energies used to convert energy differences to 0 K enthalpy differences. Finally the path connecting each TS to the nearest minimum on each side of the TS was computed using the Gonzalez-Schlegel second order intrinsic reaction

coordinate path (IRC) method [17] with a step size of 0.3 $amu^{1/2}$-bohr for the first TS and 0.05 $amu^{1/2}$-bohr for the second TS. For the reaction path involving $TiCp_2$ selected stationary points were reoptimized as noted above. The TS was also confirmed by computing its hessian. Due to the high computational cost of the large $TiCp_2$ reaction system, hessians, and thus ZPE corrections, were not performed on the minima.

III. Results and Discussion

Figure 2 shows the energy profile of each of the reaction systems. The zero of energy on the curve for each reaction is the sum of the reactant energies (structures **a**, **b**, and **c**) for that reaction. The MP2 structures at each unique stationary point are given separately for the catalysts TiH_2, $TiCl_2$, and $TiCp_2$ in Figures 3, 4 and 5 respectively [18]. Animations of the IRCs shown in Fig. 2 are available as supplementary material [19]. Structures that are identical for all three reactions, structures **b**, **c**, and **o**, are shown only in Fig. 3. The MP2 total energies and the MP2 vibrational zero point energy (ZPE) corrections for each geometry point marked in Figs. 2-5 are available as supplementary material. MP2 ZPE corrected energies are listed relative to the zero of energy in Table 1. Note that MP2 ZPE corrections were not performed on many of the geometeries involving $TiCp_2$ due to the high computational cost.

It is important to note that all points on the energy plot in Fig. 2 lie below the energy of the reactants, in contrast to the large barrier in the uncatalyzed reactions. In addition, all three reaction systems in the current study exhibit quite similar stationary points, although there are some differences in the energetics of the reactions. The following text will lay out

the reaction in detail for $TiCl_2$, shown in Fig. 4, noting the differences in the TiH_2 and $TiCp_2$ reactions where appropriate.

There are two possibilities for the first step of the reaction, both of which are barrierless processes. The first, and more exothermic, is to add the catalyst across the ethylene double bond to form the three membered ring compound shown in Fig. 4d. This process is downhill in energy by 59.6 kcal/mol at the ZPE corrected MP2 level of theory. Note that, based on the large exothermicity and the large (0.15Å) increase in the CC bond length, structure **d** is a three-membered ring, not a π complex. Silane will then add to form the complex depicted in Fig. 4e. This second barrierless addition is downhill by 5.3 kcal/mol.

In our previous work (reference 11) we found an alternative pathway with TiH_2 adding to SiH_4 first. This addition resulted in a Ti insertion into an Si-H bond. However, the subsequent addition of C_2H_4 resulted in a barrierless rearrangement to structure **e**. In the present study using $SiCl_3H$ instead of SiH_4, neither $TiCl_2$ nor TiH_2 inserted into the Si-H bond. Instead, a simple complex is formed as shown in structure **d'**. Since there is no insertion, the addition of $SiCl_3H$ to $TiCl_2$ (or TiH_2) is much less exothermic than the addition of SiH_4 (10.6 kcal/mol versus 31.1kcal/mol). Just as in the previous work, once C_2H_4 is added the complex spontaneously rearranges to structure **e**. Since neither $TiCl_2$ nor TiH_2 inserted into the Si-H bond, it is unlikely that $TiCp_2$ would insert. Thus an optimization of a structure similar to **d'** using $TiCp_2$ was not attempted.

The net result after the two reactants and the catalyst have been added together is the formation of compound **e**, with no intervening barrier in either of the two possible routes.

The overall exothermicity to this point is more than 60 kcal/mol in all the systems studied driven predominately by the addition of the catalyst to ethylene. This very large drop in energy drives the entire reaction path down in energy. In fact, the reaction path is forced down enough that all subsequent points are below the reactants in energy.

Recall that the final desired product is ethyltrichlorosilane. So, starting from compound 4e, the silyl group needs to migrate to the nearest (α) carbon, and a hydrogen needs to be transferred to the adjacent (β) carbon, with the ultimate removal of the catalyst. Therefore, the next step in the reaction is to transfer a H from the complexed silane to the Ti and to attach the Si to the α C.

In our previous study, it was found that the system went through a small barrier to structure **g** which has the silane rotated such that there are two bridging hydrogens. However, the small barrier went to zero when ZPE corrections were added. For $SiCl_3H$, a similar minimum would have one bridging hydrogen and one bridging chlorine. However, no such minimum could be found for any of the catalysts studied. Thus, it is likely that the TS represented by structure **f** in the previous study simply does not exist for the reaction involving $SiCl_3H$.

In the model system studied previously a small (5.3 kcal/mol) barrier was found which connected structure **e** (through structure **g**) to structure **i**. For reaction **III** a similar TS, shown in Figure 4h, has been located 6.3 kcal/mol in energy above structure **e**. The IRC's from the TS are shown in Fig. 2. They illustrate the extreme flatness of the surface in the TS region, which is the reason the correct TS is so difficult to locate. The IRC leading towards the products for structure **h** stops after 11 steps and after lowering the energy by 0.006

kcal/mol due to a very small gradient. Optimizations from the end of the IRC show that there might be two possible results of this step. The first is the desired structure **i** for which a reaction path has been constructed using a constrained optimization and an internal coordinate based linear least motion path. The highest point on this path connecting structure **e** with structure **i** is less than 6.5 kcal/mol (without ZPE correction). The second outcome shows that the transfer of a Cl from the Si to the Ti may be a competitive path to the desired reaction path. However, the mechanism for this step reduces the bond angle between the two-Cl's on Ti to 106° during the reaction. Since Cp rings are much larger, steric hindrance is likely to prevent this process in the mechanism of this step for the real catalyst, $TiCp_2$.

After going through the first TS the reaction proceeds to form the four-membered ring shown in Figure 4i, which is 0.8 kcal/mol above structure 4e. The four membered ring can be opened by breaking the Si-Ti bond to give structure 4k. The barrier for the model all-hydrogen system (**I**) is 3.2 kcal/mol. Structure 4k lies 6.4 kcal/mol lower in energy than structure 4j and is the global minimum on the reaction surface at 70.5 kcal/mol below the reactants in energy. Since 4k is lower in energy than 4i, it is likely the barrier is also lower than for the model reaction (**I**). Because it is unlikely that structure **j** is qualitatively important to the overall reaction path, no attempt was made to locate it for the more complex reactions.

The final steps in the mechanism are to regenerate the catalyst by transfer of a hydrogen from Ti to C and then to eliminate the catalyst. The transition state for this process is shown in structure 4l; the associated barrier height is 26.0 kcal/mol. Note, however, that this TS is still 44.5 kcal/mol lower in energy than the initial reactants.

In our previous work, the IRC for system **I** from this TS toward the products was found to proceed through a structure **m**, in which the catalyst is complexed to the ethylsilane product, then to structure **n**, in which the titanium inserts into one of the Si-H bonds. Since, for $SiCl_3H$ there are no remaining Si-H bonds, the IRC leads to the expected structure **n** with no corresponding intermediate insertion product. The complex 4n lies 7.3 kcal/mol below the TS and 14.6 kcal/mol below the seperated products in energy. From structure **n** it is quite straightforward to remove the catalyst from the complex.

Once the catalyst is removed, the process is complete with ethyltrichlorosilane as the product. The overall process is exothermic by 37.2 kcal/mol at the ZPE corrected MP2 level of theory. This compares with the value of 37.1 kcal/mol computed by Day and Gordon at the MP2/6-311G(d,p) level of theory.

The driving force for the entire reaction comes in the first two steps with the formation of the compound shown in structure **e**, which is 64.9 kcal/mol below the reactants in energy. The reasons this structure is so stable are illustrated by the first two steps in the reaction. In the first step the electron deficient catalyst adds to the ethylene across the π bond in much the same manner as the addition of CH_2 to ethylene to form cyclopropane. The second step is much less exothermic and is driven mostly by the electrostatic attraction between the positively charged titanium (+0.47) and the negative hydrogen (-0.09) on the silicon.

IV. Conclusions

The results presented here clearly show that divalent titanium is an effective catalyst for the hydrosilation reaction. All four reaction systems studied show qualitatively the same behavior. The most significant change made was the substitution of 3 chlorines for 3 hydrogens. This substitution prevented the unwanted insertion of titanium into Si-H bonds, but did not alter the qualitative aspects of the reaction path. The effect of the substituents on the titanium is smaller. While there is some change in the energetics, particularly with $TiCp_2$, there are not any qualitative changes in the main reaction path.

The overall catalyzed reaction has no net barrier, because of the very stable cyclic $TiX_2CH_2CH_2$ intermediate. However, the energy profile of the multistep process (Fig. 2) does offer the possibility of finding some of the intermediate structures if the process was carried out at low temperature.

Acknowledgements

The work described in this paper was supported by grants from the National Science Foundation (CHE-9633480) and the Air Force Office for Scientific Research (F49-620-95-1-0073). The computations were performed in part on computers provided by Iowa State University and on the CEWES T3E via the Department of Defense Grand Challenge program.

References

1. Barton TJ, Boudjouck P, "Organosilicon chemistry - a brief overview", in Advances in Silicon-Based Polymer Science, Advances in Chemistry Series No. 224, Ziegler J.; Fearon, F. W. G. (eds) American Chemical Society, Washington, DC. 1990, 3-46
2. Speier JL (1979) Adv. Organomet. Chem. 17: 407
3. Day PN, Gordon MS (1995) Theor Chim Acta 91: 83
4. Harrod JF, Ziegler T, Tschinke V (1990) Organometallics 9: 897
5. Koga N, Morokuma K (1988) J Am Chem Soc 110: 108
6. Yoshida T, Koga N, Morokuma K (1995) Organometallics 14: 746
7. Jensen VR, Børve KJ, Ystenes M (1995) J Am Chem Soc 117: 4109
8. Sakai S (1994) J Phys Chem 98: 12053
9. Kawamura-Kuribayashi H, Koga N, Morokuma K (1992) J Am Chem Soc 114: 2359
10. Stahl M, Pidun U, Frenking G (1997) Angew Chem Int Ed Engl 36: 2234
11. Bode BM, Day PN, Gordon MS (1998) J Am Chem Soc 120: 1552
12. Bode BM, Raaii F, Gordon MS Catalysis of the hydrosilation and bis-silylation reactions. In *Transition State Modeling of Catalysis in Computational Chemistry*; ACS Symposium Series, in press.
13. (a) For C, Si, Cl: Stevens WJ, Basch H, Krauss M (1984) J Chem Phys 81: 6026; (b) For Ti: Stevens WJ, Basch H, Krauss M, Jasien P (1992) Can J Chem 70: 612
14. The exponents used were: C $\zeta_d = 0.590$, Si $\zeta_d = 0.364$, Cl $\zeta_d = 0.566$

15. (a) For H: DitchField R, Hehre WJ, Pople JA (1971) J Chem Phys 54: 724; (b) For C: Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56: 2257; (c) The polarization exponents, H ζ_p = 1.1 and C ζ_d = 0.8, are given by: Hariharan PC, Pople JA (1973) Theoret Chim Acta 28: 213

16. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA Jr The general atomic and molecular electronic structure system (1993) J Comp Chem 14: 1347

17. (a) Gonzalez C, Schlegel HB (1990) J Phys Chem 94: 5523; (b) Gonzalez C, Schlegel HB (1991) J Chem Phys 95: 5853

18. The full set of cartesian coordinates, MP2 total energies, and MP2 ZPE corrections for each structure are available as an Appendix and on the WWW site: http://www.msg.ameslab.gov/

19. Animations of the IRC's are available on the WWW site: http://www.msg.ameslab.gov/

Table 1: MP2 relative energies (kcal/mol). The first three reactions include an MP2 ZPE correction. Those involving $TiCp_2$ are not ZPE corrected.

Geometry point	**I** $TiH_2 + SiH_4 + C_2H_4$	**II** $TiH_2 + SiCl_3H + C_2H_4$	**III** $TiCl_2 + SiCl_3H + C_2H_4$	**IV** $TiCp_2 + SiCl_3H + C_2H_4$
a + b + c (reactants)	0	0	0	0
d' + b	-31.1	-8.3	-10.6	
d + c	-61.9	-61.8	-59.6	-47.2
e	-68.4	-66.2	-64.9	-61.2
f	-66.6			
g	-66.6			
h	-61.3		-58.6	
i	-67.2	-70.0	-64.1	
j	-64.0			
k	-65.4	-74.3	-70.5	-72.1
l	-33.5	-42.3	-44.5	-42.0
m	-39.0			
n	-59.2	-46.9	-51.8	-51.4
o + a (products)	-28.0	-37.2	-37.2	-41.6

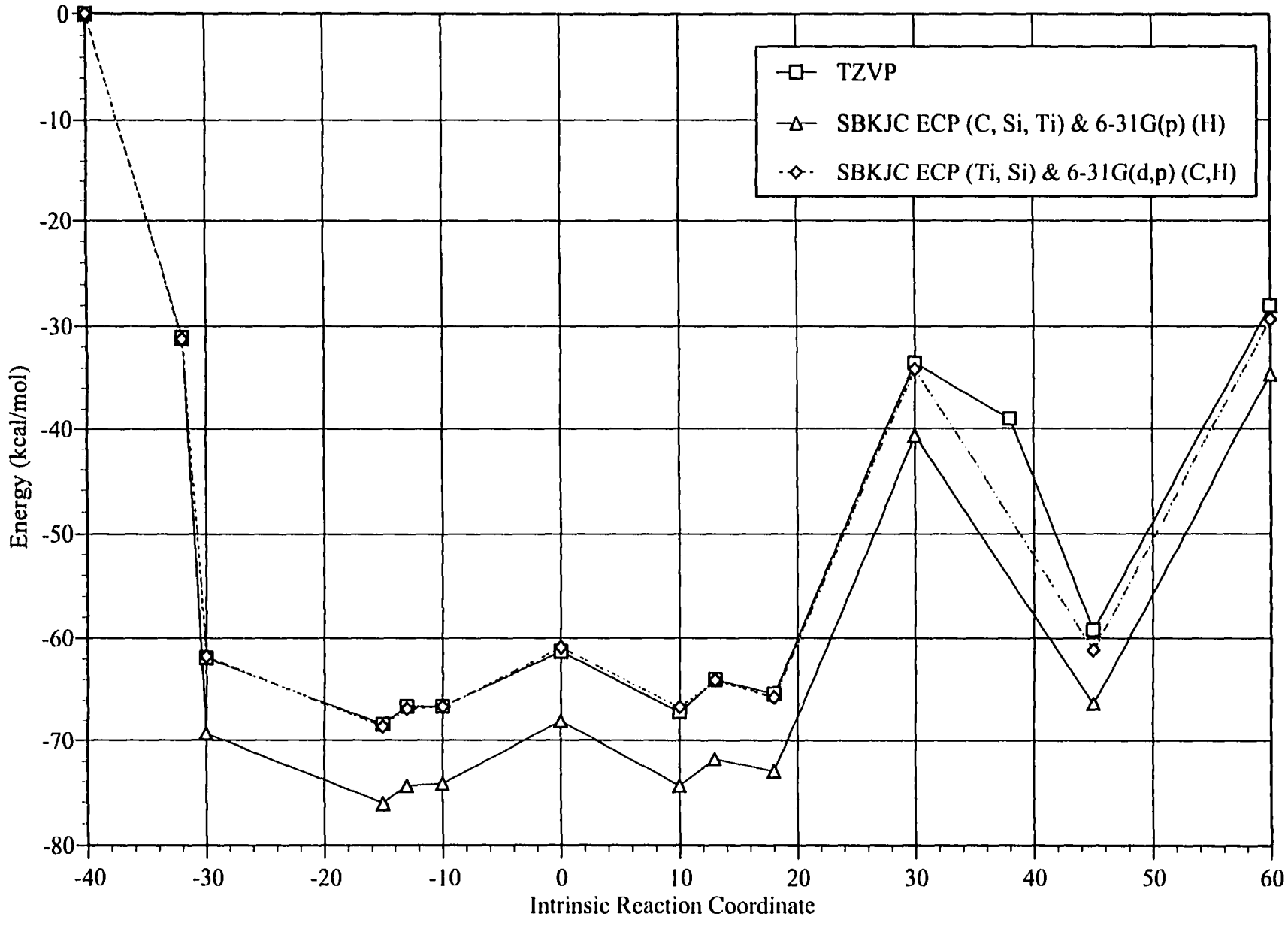


Figure 1: Comparison of TZVP versus SBK ECP basis sets

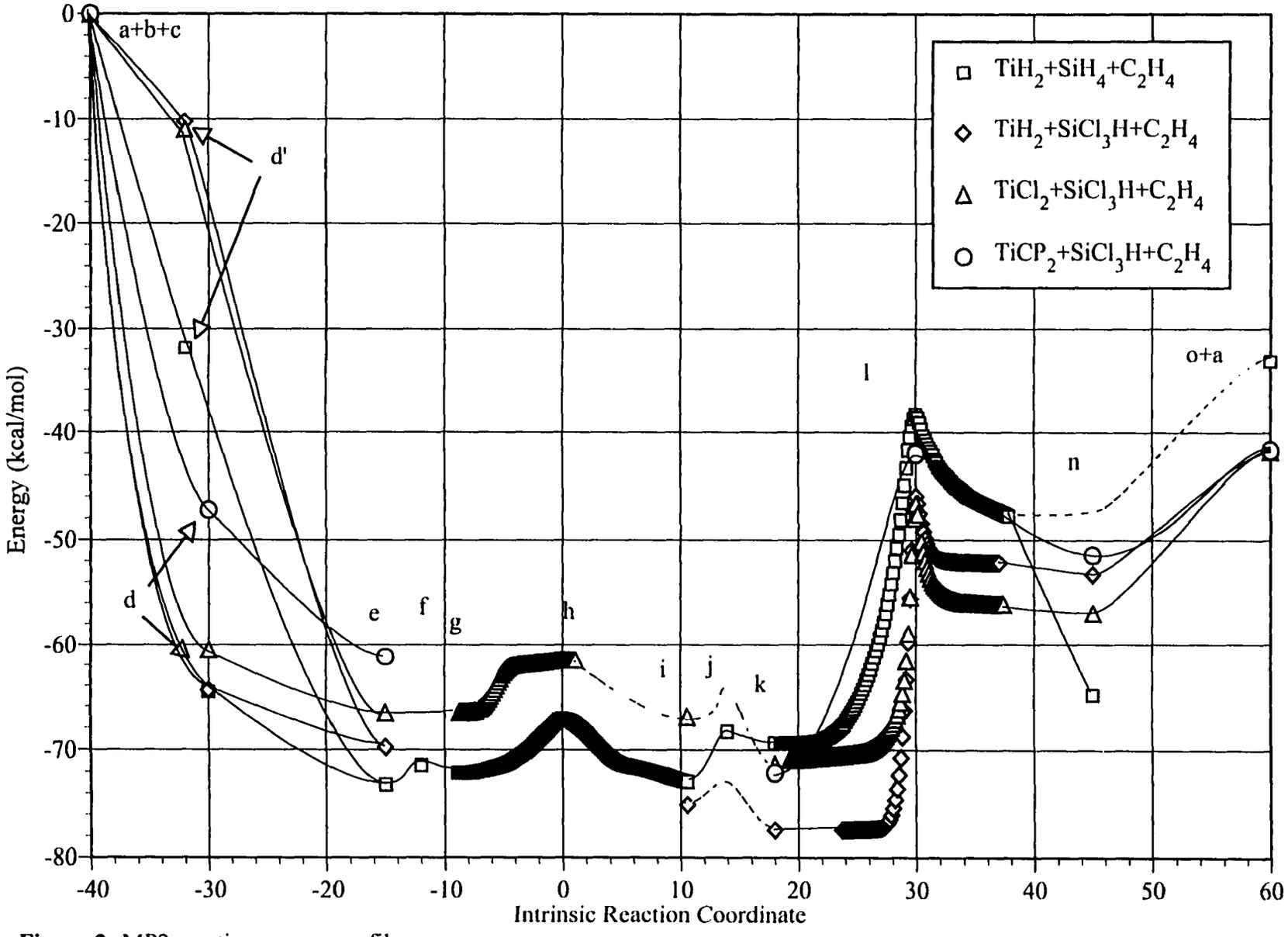


Figure 2: MP2 reaction energy profiles

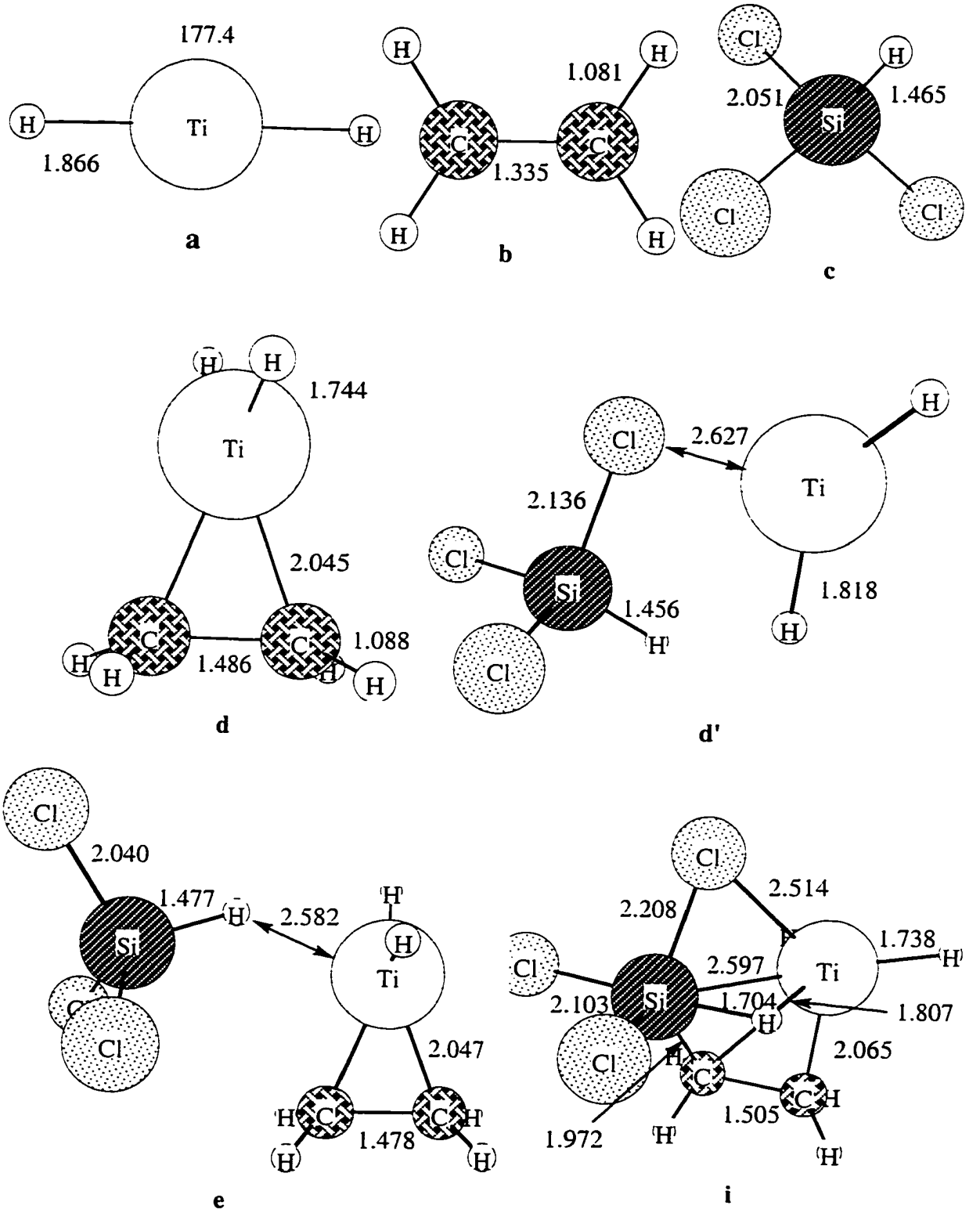
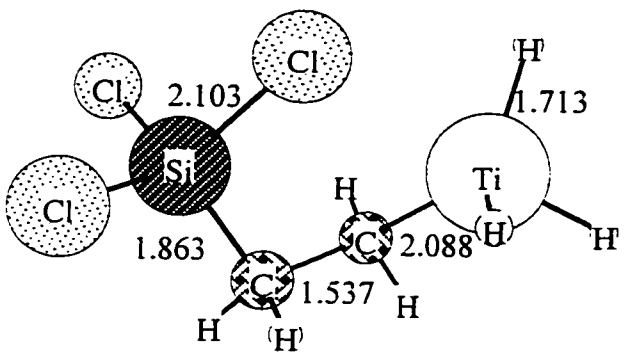
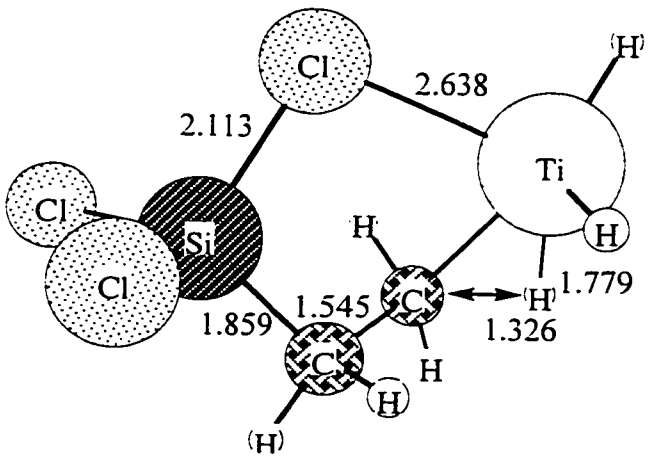


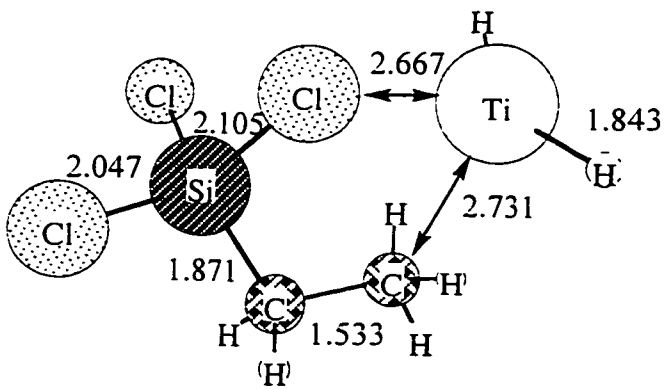
Figure 3: MP2 structures along the minimum energy reaction path



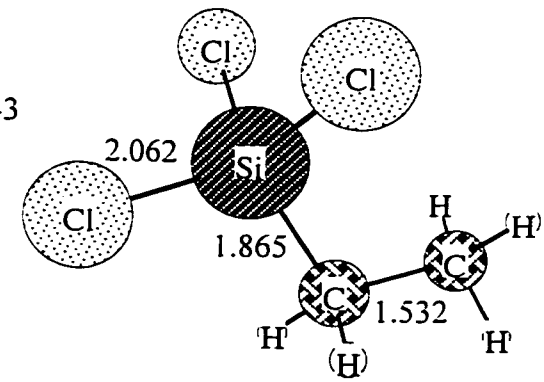
k



l



n



o

Figure 3: continued

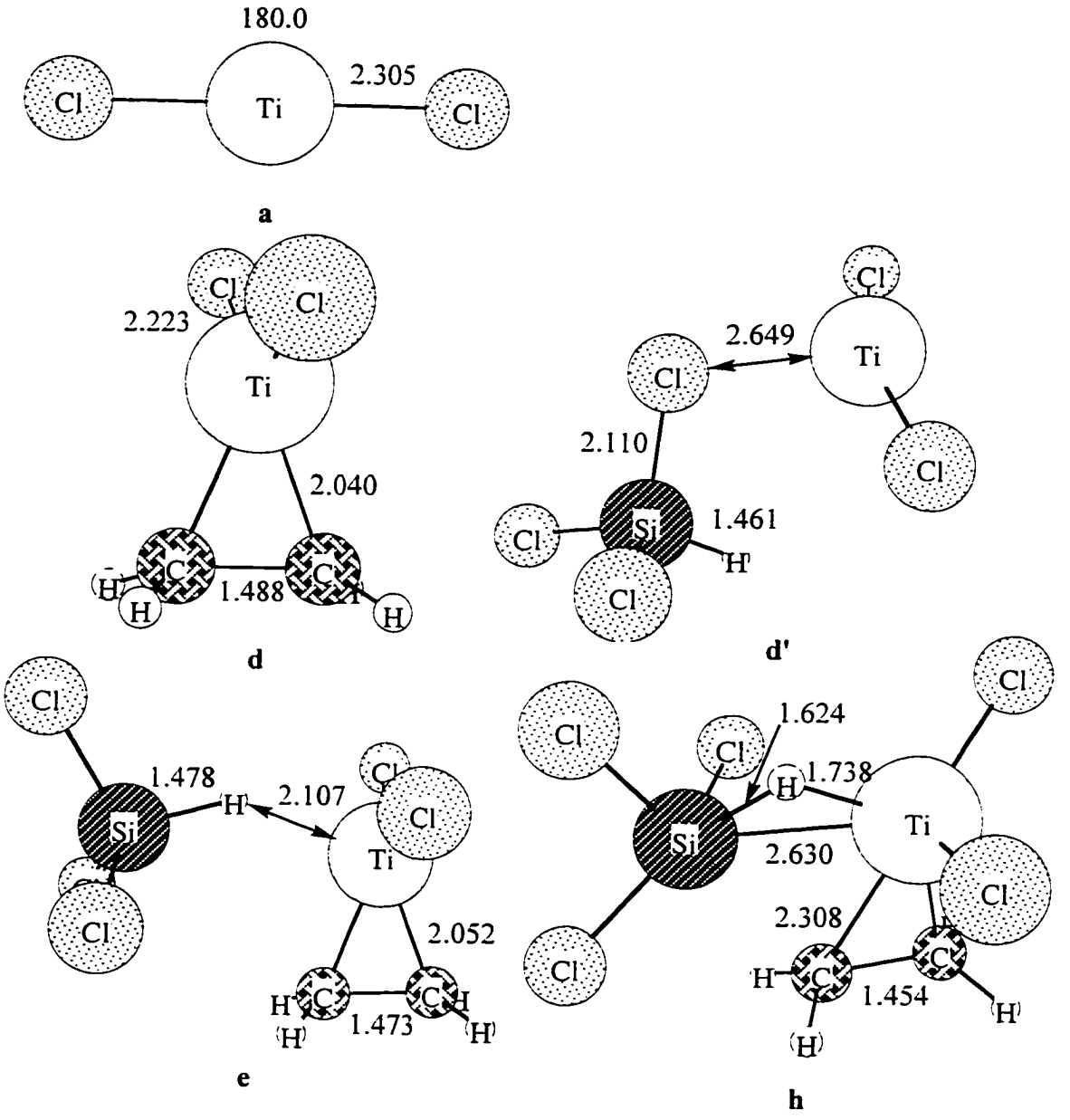
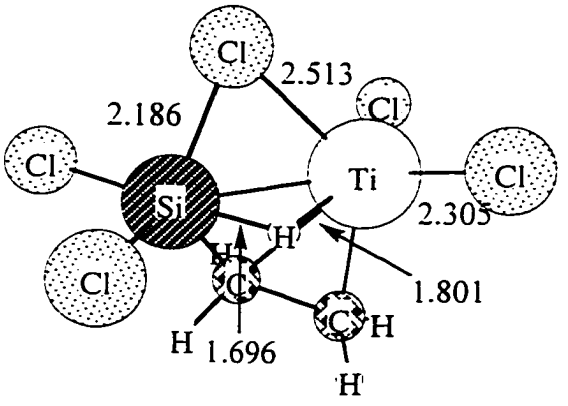
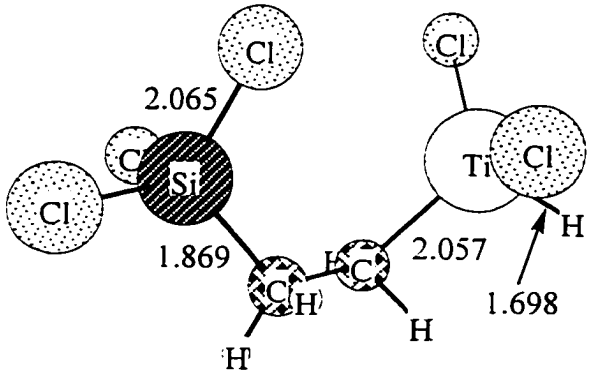


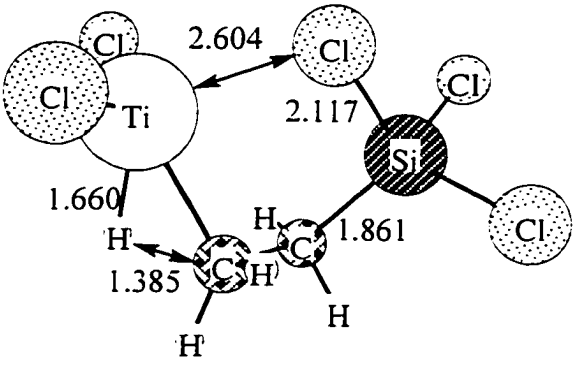
Figure 4: MP2 structures along the minimum energy reaction path using $TiCl_2$ as the catalyst



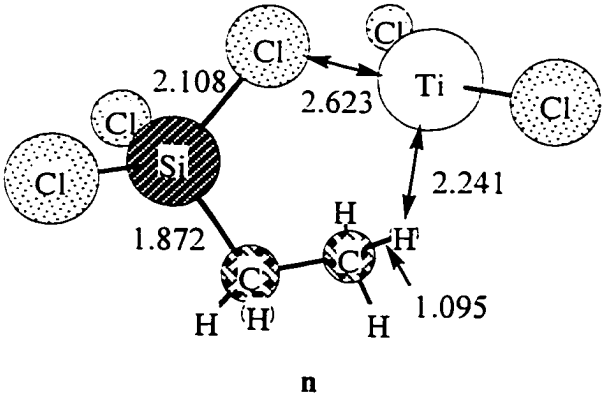
i



k

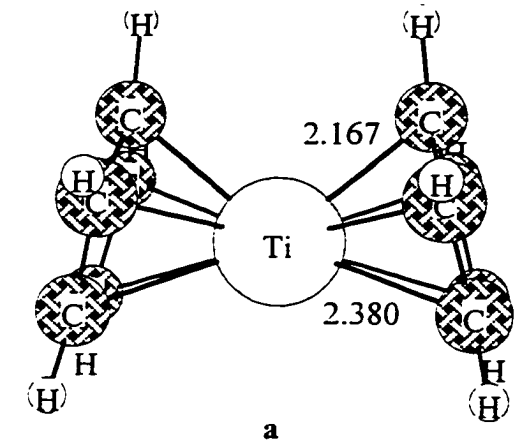


l

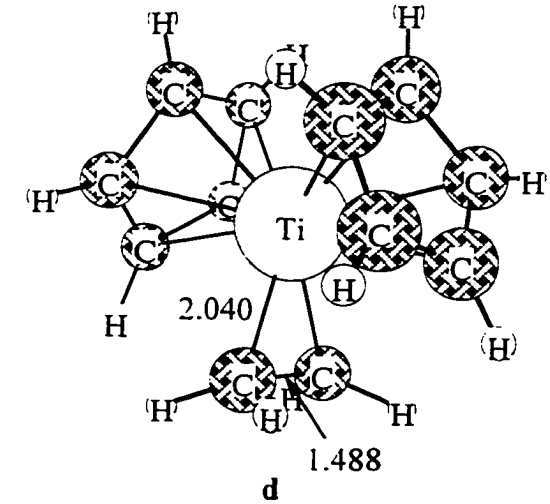


n

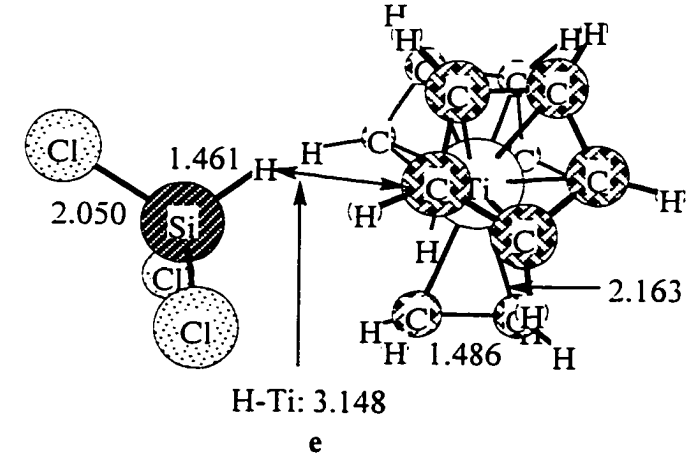
Figure 4: Continued



a

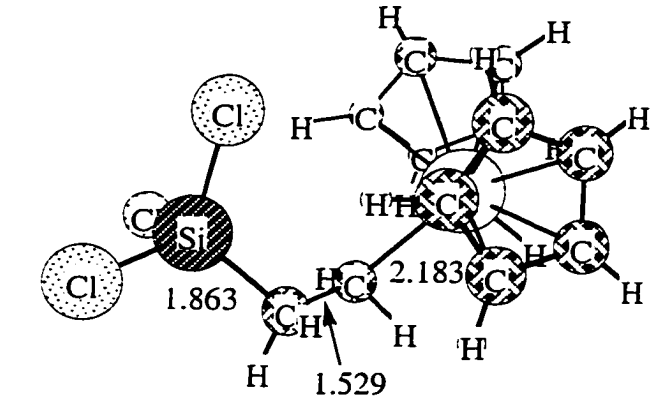


d

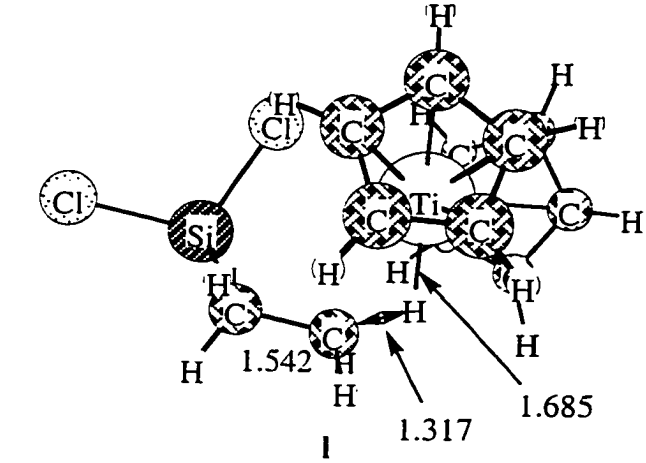


e

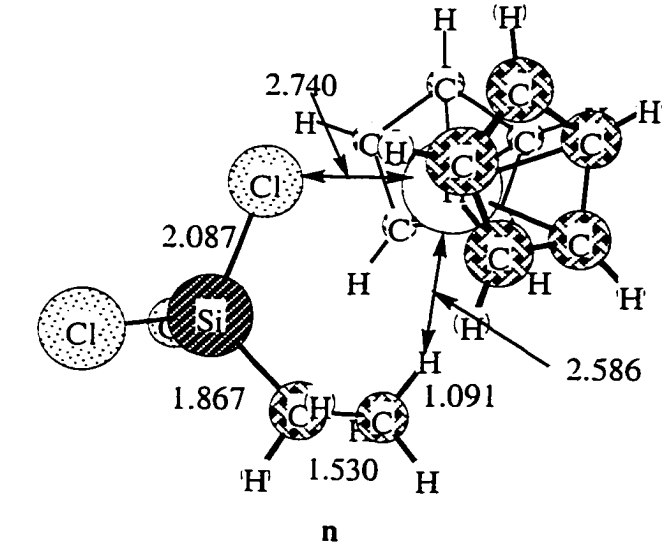
Figure 5: MP2 structures along the minimum energy path using $TiCP_2$ as the catalyst



k



l



n

Figure 5: Continued

Appendix:

MP2 Total Energies in Hartrees for each reaction system:

Geometry point	II $TiH_2 + SiCl_3H + C_2H_4$	III $TiCl_2 + SiCl_3H + C_2H_4$	IV $TiCp_2 + SiCl_3H + C_2H_4$
a	-58.582174	-87.281162	-443.436587
b	-78.317282	-78.317282	-78.317282
c	-49.032702	-49.032702	-49.032702
d'	-107.631227	-136.331341	
d	-137.002133	-165.694824	-521.829078
e	-186.043342	-214.737267	-570.884121
f			
g			
h		-214.729192	
i	-186.051765	-214.737830	
j			
k	-186.055543	-214.744744	-570.901434
l	-186.005280	-214.705339	-570.853483
m			
n	-186.017059	-214.711832	-570.868503
o	-127.416311	-127.416311	-127.416311

MP2 zero point energy (ZPE) corrections (in Hartrees):

Geometry point	**II** $TiH_2 + SiCl_3H + C_2H_4$	**III** $TiCl_2 + SiCl_3H + C_2H_4$	**IV** $TiCp_2 + SiCl_3H + C_2H_4$
a	0.007098	0.002153	
b	0.052344	0.052344	0.052344
c	0.014483	0.014483	0.014483
d'	0.024627	0.017300	
d	0.063675	0.055931	
e	0.079623	0.071703	
f			
g			
h		0.073700	
i	0.082053	0.073437	
j			
k	0.078838	0.070207	
l	0.079660	0.072182	0.242016
m			
n	0.084093	0.077192	
o	0.073828	0.073828	0.073828

Cartesian Coordinates for each stationary point along the reaction **II**: $TiH_2 + SiH_4 + C_2H_4$

Point **a**:

Ti	22.0	0.00534	0.00000	0.00050
H	1.0	-0.12771	0.00000	1.86208
H	1.0	0.22164	0.00000	-1.85324
H	1.0	0.22164	0.00000	-1.85324

Point **b**:

C	6.0	0.00000	0.00000	0.66763
C	6.0	0.00000	0.00000	-0.66763
H	1.0	0.00000	0.92079	1.23332
H	1.0	0.00000	-0.92080	1.23329
H	1.0	0.00000	0.92079	-1.23332

H	1.0	0.00000	-0.92080	-1.23329

Point **c**:

Si	14.0	-0.01229	0.23776	0.10182
H	1.0	0.83374	1.10056	0.92945
Cl	17.0	-1.13657	-0.97774	1.31296
Cl	17.0	-1.25299	1.39974	-1.04656
Cl	17.0	1.17243	-0.91238	-1.11553

Point **d**:

Ti	22.0	0.00006	0.75244	-0.00003
C	6.0	0.74254	-1.15250	0.00010
C	6.0	-0.74310	-1.15239	-0.00002
H	1.0	0.00215	1.58254	1.53396
H	1.0	0.00036	1.58374	-1.53351
H	1.0	1.24605	-1.46977	-0.91057
H	1.0	1.24593	-1.46927	0.91101
H	1.0	-1.24654	-1.47025	0.91046
H	1.0	-1.24616	-1.46975	-0.91092
H	1.0	-1.24616	-1.46975	-0.91092

Point **d'**:

Si	14.0	-0.06888	-1.45594	1.62863
H	1.0	0.90471	-1.98251	2.57435
Cl	17.0	-1.27516	-2.94783	0.91955
Cl	17.0	-1.18898	0.08280	2.33563

Cl	17.0	0.97906	-0.77371	-0.10351
Ti	22.0	2.86347	0.70873	0.96967
H	1.0	2.26279	0.29541	2.63533
H	1.0	4.21994	1.69906	0.27382

Point **e**:

C	6.0	-3.07200	-0.20051	0.10894
C	6.0	-1.99062	0.43682	0.88915
Ti	22.0	-1.27970	-1.09680	-0.30909
Si	14.0	1.45240	-0.14649	1.47620
H	1.0	-3.78882	-0.81859	0.64568
H	1.0	-3.50432	0.36166	-0.71662
H	1.0	-1.68711	1.44016	0.59816
H	1.0	-1.98670	0.27849	1.96464
H	1.0	-1.03137	-2.66197	0.42125
H	1.0	-0.57351	-0.62287	-1.83183
H	1.0	0.71234	-0.75906	0.35403
Cl	17.0	1.39030	1.88591	1.31201
Cl	17.0	3.37257	-0.81556	1.26777
Cl	17.0	0.67809	-0.76756	3.25729

Point **i**:

C	6.0	-0.11724	-0.17349	-0.03912
C	6.0	-0.04897	0.18845	1.42051
Ti	22.0	1.93840	0.02255	-0.02548

Si	14.0	1.48128	-0.45387	2.48602
H	1.0	-0.51771	-1.15509	-0.27375
H	1.0	-0.50755	0.61211	-0.68459
H	1.0	-0.70549	-0.47571	2.00055
H	1.0	-0.28343	1.21832	1.66312
H	1.0	1.74677	-1.38923	1.08623
H	1.0	2.49748	-0.37402	-1.62265
H	1.0	1.68370	1.69274	0.14608
Cl	17.0	1.39189	-2.29278	3.43578
Cl	17.0	1.22372	0.82421	4.13596
Cl	17.0	3.48764	0.28704	1.93692

Point **k**:

C	6.0	-1.30965	1.71473	0.52557
C	6.0	-0.21767	1.70157	1.60763
Ti	22.0	-2.03525	-0.18007	0.03471
Si	14.0	1.33854	0.93571	0.92822
H	1.0	-0.98678	2.23298	-0.38479
H	1.0	-2.21081	2.22470	0.88942
H	1.0	0.02184	2.70301	1.97865
H	1.0	-0.51920	1.10870	2.47596
H	1.0	-1.94590	-0.68589	-1.59996
H	1.0	-3.72740	0.06856	0.07330
H	1.0	-1.91402	-1.30265	1.32965

Cl	17.0	0.71173	-0.75243	-0.15810
Cl	17.0	2.66826	0.31307	2.35665
Cl	17.0	2.28613	2.17659	-0.40310

Point **l**:

C	6.0	-1.31272	1.86278	0.34717
C	6.0	-0.25699	1.79850	1.47311
Ti	22.0	-1.93316	-0.35288	0.19748
Si	14.0	1.30602	0.98746	0.87682
H	1.0	-0.88488	1.92795	-0.66077
H	1.0	-1.87818	2.78978	0.46542
H	1.0	0.00306	2.79240	1.84717
H	1.0	-0.63258	1.23365	2.33371
H	1.0	-2.73058	-0.87342	-1.30169
H	1.0	-2.45708	1.22731	0.55599
H	1.0	-1.96026	-1.17915	1.77292
Cl	17.0	0.64124	-0.63243	-0.30603
Cl	17.0	2.49004	0.22544	2.35831
Cl	17.0	2.41421	2.21025	-0.33850

Point **n**:

C	6.0	-1.38645	1.86163	1.23182
C	6.0	-0.01726	1.90482	1.92068
Ti	22.0	-1.88328	-0.19186	-0.49871
Si	14.0	1.32123	0.93625	1.04343

H	1.0	-1.32421	2.02769	0.14852
H	1.0	-2.03955	2.64630	1.60954
H	1.0	0.35056	2.93152	1.97731
H	1.0	-0.09253	1.53548	2.94444
H	1.0	-1.71295	0.59498	-2.15684
H	1.0	-1.91088	0.92792	1.46639
H	1.0	-3.03941	-1.15906	0.55987
Cl	17.0	0.52623	-0.90280	0.39764
Cl	17.0	2.90746	0.52062	2.26878
Cl	17.0	1.95208	1.92404	-0.63375

Point **o**:

C	6.0	-1.93251	-0.27912	0.20303
C	6.0	-0.92943	0.77351	0.68441
Si	14.0	0.84317	0.19483	0.63693
Cl	17.0	2.10593	1.68248	1.30211
Cl	17.0	1.09415	-1.46320	1.83759
Cl	17.0	1.38576	-0.31319	-1.28689
H	1.0	-2.94731	0.11191	0.25124
H	1.0	-1.73527	-0.56964	-0.82729
H	1.0	-1.89030	-1.17509	0.81973
H	1.0	-1.13945	1.07234	1.71337
H	1.0	-0.98596	1.67627	0.07285

Cartesian Coordinates for each stationary point along the reaction **III**: $TiCl_2 + SiCl_3H + C_2H_4$

Point **a**:

Ti	22.0	0.00000	0.00000	0.00000
Cl	17.0	0.00000	0.00000	-2.30481
Cl	17.0	0.00000	0.00000	2.30481

Points **b** and **c**: same as reaction **II**

Point **d**:

Ti	22.0	-0.35268	0.55413	0.11218
C	6.0	0.59426	-1.25311	0.08911
C	6.0	-0.87976	-1.41157	-0.03471
Cl	17.0	-0.61268	1.37473	2.16199
Cl	17.0	-0.29905	1.59111	-1.85368
H	1.0	1.19683	-1.46253	-0.79064
H	1.0	1.05469	-1.55962	1.02451
H	1.0	-1.41296	-1.82534	0.81703
H	1.0	-1.27095	-1.72750	-0.99809

Point **d'**:

Si	14.0	-0.12410	-1.61285	1.55149
H	1.0	1.01897	-1.75245	2.45042
Cl	17.0	-0.63914	-3.40625	0.72606
Cl	17.0	-1.71875	-0.74067	2.46205
Cl	17.0	0.48816	-0.36144	-0.03363
Ti	22.0	2.65739	0.82309	0.92011

Cl	17.0	1.72388	1.64867	2.86857
Cl	17.0	4.36771	-0.03160	-0.36107

Point **e**:

C	6.0	-2.36859	0.05988	-0.33220
C	6.0	-1.39767	0.56628	0.65366
Ti	22.0	-0.48768	-0.63692	-0.76586
Si	14.0	2.05543	0.27213	1.32990
H	1.0	-3.07011	-0.70579	-0.01283
H	1.0	-2.77072	0.75826	-1.06081
H	1.0	-1.14647	1.62279	0.61390
H	1.0	-1.44998	0.16571	1.66240
Cl	17.0	-0.34928	-2.79453	-0.23783
Cl	17.0	0.31916	0.44117	-2.53775
H	1.0	1.37712	-0.34758	0.17210
Cl	17.0	1.84486	2.29685	1.23775
Cl	17.0	4.01704	-0.25617	1.13254
Cl	17.0	1.28037	-0.47946	3.05774

Point **h**:

C	6.0	0.05727	0.28622	-0.00026
C	6.0	0.02531	-0.08685	1.40503
Ti	22.0	2.07181	-0.08333	0.33742
Si	14.0	1.80199	0.08318	2.94774
H	1.0	-0.32574	-0.41092	-0.73806

H	1.0	-0.09592	1.33613	-0.23870
H	1.0	-0.32238	-1.09084	1.62889
H	1.0	-0.39663	0.66444	2.05820
H	1.0	2.82176	-0.52696	1.84090
Cl	17.0	2.24247	-2.15315	-0.34083
Cl	17.0	3.31144	1.26423	-0.87918
Cl	17.0	3.43068	-0.30411	4.21476
Cl	17.0	0.30506	-0.27064	4.35789
Cl	17.0	2.03515	2.09938	2.38141

Point i:

C	6.0	-0.09556	0.44339	-0.09407
C	6.0	0.02572	0.65075	1.38392
Ti	22.0	1.98593	0.57074	-0.01790
Si	14.0	1.44441	-0.36820	2.33078
H	1.0	-0.51590	-0.50617	-0.40868
H	1.0	-0.47534	1.29322	-0.65651
H	1.0	-0.69532	0.02001	1.92456
H	1.0	-0.06199	1.67157	1.74388
H	1.0	1.64764	-1.01125	0.77415
Cl	17.0	2.82401	-0.38010	-1.80273
Cl	17.0	2.14744	2.79120	-0.04783
Cl	17.0	1.10512	-2.33336	2.87367
Cl	17.0	1.22900	0.58636	4.18796

Cl	17.0	3.50281	0.26847	1.96306

Point **k**:

C	6.0	-1.01178	0.79482	0.57673
C	6.0	0.20754	0.81675	1.49283
Ti	22.0	-1.89040	-0.90022	-0.18861
Si	14.0	1.75859	0.27671	0.60044
H	1.0	-0.92811	1.49174	-0.26491
H	1.0	-1.93489	1.05499	1.12994
H	1.0	0.38741	1.81790	1.89750
H	1.0	0.07352	0.14327	2.34243
Cl	17.0	-1.46207	-1.20145	-2.31707
H	1.0	-3.51770	-0.41720	-0.21022
Cl	17.0	-1.88165	-2.47109	1.34215
Cl	17.0	1.50510	-1.66363	-0.06051
Cl	17.0	3.38944	0.34764	1.84840
Cl	17.0	2.09681	1.47850	-1.03361

Point **l**:

C	6.0	-1.02175	1.06096	-0.00541
C	6.0	0.03785	1.18747	1.11145
Ti	22.0	-1.40476	-1.17583	-0.25220
Si	14.0	1.67053	0.47432	0.57303
H	1.0	-0.61000	1.22059	-1.00893
H	1.0	-1.73560	1.87765	0.12897

H	1.0	0.21298	2.23419	1.37796
H	1.0	-0.28664	0.68003	2.02247
Cl	17.0	-2.21797	-1.69496	-2.28322
H	1.0	-2.14276	0.26022	0.13275
Cl	17.0	-1.45993	-2.24628	1.73193
Cl	17.0	1.16981	-1.18650	-0.63990
Cl	17.0	2.85471	-0.21008	2.09003
Cl	17.0	2.72877	1.76578	-0.61521

Point **n**:

C	6.0	-1.08658	0.97219	0.98847
C	6.0	0.32023	1.34751	1.46932
Ti	22.0	-1.39097	-1.42504	-0.24944
Si	14.0	1.71078	0.46764	0.57771
H	1.0	-1.16744	0.94585	-0.10489
H	1.0	-1.82343	1.70335	1.31619
H	1.0	0.50236	2.41375	1.31900
H	1.0	0.42211	1.15311	2.53792
Cl	17.0	-1.44333	-0.78033	-2.46142
H	1.0	-1.40977	0.03322	1.45205
Cl	17.0	-2.27215	-2.66119	1.49039
Cl	17.0	1.17159	-1.55045	0.29388
Cl	17.0	3.44890	0.47767	1.65692
Cl	17.0	2.01335	1.26518	-1.27897

Point **o**: Same as reaction **II**

Cartesian Coordinates for each stationary point along the reaction **IV**: $TiCp_2$+ $SiCl_3H$ + C_2H_4

Point **a**:

Ti	22.0	-0.21757	0.00009	0.00000
C	6.0	0.36857	1.16509	1.83514
C	6.0	-0.94326	0.70601	2.15400
C	6.0	-0.94323	-0.70598	2.15403
C	6.0	0.36861	-1.16502	1.83521
C	6.0	1.20367	-0.00004	-1.63631
C	6.0	0.36859	1.16504	-1.83511
C	6.0	-0.94326	0.70599	-2.15395
C	6.0	-0.94322	-0.70598	-2.15404
C	6.0	0.36863	-1.16503	-1.83527
C	6.0	1.20364	-0.00002	1.63628
H	1.0	0.69554	2.19303	1.83629
H	1.0	-1.80631	1.33202	2.33054
H	1.0	-1.80627	-1.33205	2.33058
H	1.0	0.69554	-2.19299	1.83635
H	1.0	2.27587	-0.00008	-1.52349
H	1.0	0.69560	2.19302	-1.83630
H	1.0	-1.80631	1.33200	-2.33044
H	1.0	-1.80625	-1.33207	-2.33057

H	1.0	0.69555	-2.19302	-1.83642
H	1.0	2.27586	-0.00001	1.52347

Points **b** and **c**: Same as reaction **II**

Point **d**:

Ti	22.0	-0.04818	-0.04008	-0.09574
C	6.0	0.08129	1.64853	1.62053
C	6.0	-0.87495	0.72520	2.13126
C	6.0	-0.24878	-0.53085	2.29898
C	6.0	1.11365	-0.41588	1.89531
C	6.0	1.76096	0.33292	-1.50561
C	6.0	0.91989	1.50848	-1.55819
C	6.0	-0.30929	1.13685	-2.17509
C	6.0	-0.25332	-0.23587	-2.50615
C	6.0	1.01265	-0.74916	-2.10237
C	6.0	1.33616	0.94610	1.46688
H	1.0	-0.08212	2.70287	1.46257
H	1.0	-1.92163	0.93181	2.30787
H	1.0	-0.73713	-1.43695	2.62308
H	1.0	1.86259	-1.18815	1.98083
H	1.0	2.80462	0.31010	-1.23098
H	1.0	1.20171	2.50726	-1.26577
H	1.0	-1.16587	1.78128	-2.31534
H	1.0	-1.05921	-0.80966	-2.93956

H	1.0	1.37690	-1.74814	-2.29187
H	1.0	2.29167	1.39374	1.23985
C	6.0	-1.11848	-1.77593	-0.14935
C	6.0	-2.08990	-0.65167	-0.23786
H	1.0	-1.03594	-2.39128	-1.04312
H	1.0	-1.19406	-2.41472	0.72528
H	1.0	-2.74557	-0.50738	0.61374
H	1.0	-2.61237	-0.51657	-1.17925

Point e:

Ti	22.0	-2.00438	0.22437	1.79736
C	6.0	-3.34441	0.58152	3.65661
C	6.0	-2.78355	-0.71125	3.82683
C	6.0	-1.38301	-0.55946	3.97732
C	6.0	-1.07208	0.81857	3.91423
C	6.0	-2.47828	2.16057	0.61569
C	6.0	-3.70225	1.43705	0.73709
C	6.0	-3.57618	0.23485	-0.00291
C	6.0	-2.28506	0.21343	-0.58166
C	6.0	-1.60705	1.40036	-0.21203
C	6.0	-2.28027	1.53303	3.70598
H	1.0	-4.39277	0.80854	3.54684
H	1.0	-3.31981	-1.64713	3.81628
H	1.0	-0.66888	-1.36061	4.07851

H	1.0	-0.08162	1.23712	3.99107
H	1.0	-2.27628	3.13417	1.03482
H	1.0	-4.57531	1.76148	1.28025
H	1.0	-4.31416	-0.54972	-0.07713
H	1.0	-1.87003	-0.58569	-1.17374
H	1.0	-0.60638	1.66515	-0.51110
H	1.0	-2.38647	2.60438	3.63666
C	6.0	-0.22429	-0.89390	1.28666
C	6.0	-1.40225	-1.79835	1.31993
H	1.0	0.21906	-0.72082	0.30987
H	1.0	0.53497	-1.05111	2.04755
H	1.0	-1.43475	-2.52907	2.12298
H	1.0	-1.72320	-2.23321	0.37821
Si	14.0	1.86788	2.20861	1.56515
H	1.0	0.42043	2.29801	1.74296
Cl	17.0	2.68630	0.98053	2.98931
Cl	17.0	2.66626	4.09433	1.77642
Cl	17.0	2.31118	1.50604	-0.30950

Point **k**:

Ti	22.0	-1.84123	-0.57754	-1.24471
C	6.0	-1.67910	-0.15563	-3.59344
C	6.0	-0.94324	0.85728	-2.92465

C	6.0	0.14681	0.25151	-2.25691
C	6.0	0.07028	-1.14747	-2.47642
C	6.0	-1.05392	-1.39450	-3.31540
C	6.0	-1.68338	-2.82790	-0.50686
C	6.0	-2.90376	-2.68540	-1.21990
C	6.0	-3.73971	-1.79576	-0.49567
C	6.0	-3.02409	-1.37886	0.65610
C	6.0	-1.75937	-2.01957	0.65191
H	1.0	-2.56156	-0.00604	-4.19350
H	1.0	-1.19721	1.90610	-2.90301
H	1.0	0.88802	0.76080	-1.66406
H	1.0	0.76923	-1.88310	-2.11144
H	1.0	-1.36788	-2.35906	-3.68168
H	1.0	-0.84653	-3.44551	-0.79330
H	1.0	-3.15898	-3.17242	-2.14778
H	1.0	-4.74259	-1.50763	-0.76132
H	1.0	-3.38255	-0.68588	1.40143
H	1.0	-0.98243	-1.91010	1.38966
C	6.0	-1.43988	1.06705	0.13410
C	6.0	-0.90869	1.03329	1.56734
Si	14.0	0.94452	0.85941	1.63492
H	1.0	-0.98166	1.90281	-0.39652
H	1.0	-2.52653	1.27367	0.15322

H	1.0	-1.12428	1.95301	2.12695
H	1.0	-1.32571	0.21375	2.15810
H	1.0	-3.32830	0.08774	-1.84009
Cl	17.0	1.62781	-0.93472	0.85621
Cl	17.0	1.61545	0.97059	3.58695
Cl	17.0	1.85806	2.36936	0.55784

Point I:

C	6.0	-1.33275	1.42932	0.57876
C	6.0	-1.03008	1.27653	2.08302
Si	14.0	0.77452	1.05831	2.44121
H	1.0	-0.79342	2.29172	0.18601
H	1.0	-2.39659	1.69761	0.52764
H	1.0	-1.37648	2.15038	2.65081
H	1.0	-1.53304	0.40637	2.51049
H	1.0	-1.55300	0.69665	-0.49333
Cl	17.0	1.42515	-0.65238	1.45034
Cl	17.0	1.14445	0.78992	4.45238
Cl	17.0	1.90541	2.64625	1.77733
Ti	22.0	-0.43245	-0.55453	-0.63049
C	6.0	-0.19913	-0.35398	-2.94663
C	6.0	-0.22327	0.92883	-2.33529
C	6.0	0.93939	1.01436	-1.49221
C	6.0	1.67040	-0.20876	-1.62990

C	6.0	0.96228	-1.04630	-2.51192
C	6.0	-1.55909	-2.33921	-1.37689
C	6.0	-2.53698	-1.43875	-0.84746
C	6.0	-2.34140	-1.40504	0.56501
C	6.0	-1.25611	-2.24145	0.89354
C	6.0	-0.74054	-2.81348	-0.29912
H	1.0	-0.94611	-0.73415	-3.62532
H	1.0	-0.91594	1.72430	-2.55680
H	1.0	1.27197	1.88684	-0.95494
H	1.0	2.60322	-0.44937	-1.14562
H	1.0	1.22753	-2.06030	-2.77142
H	1.0	-1.49615	-2.67432	-2.39857
H	1.0	-3.34808	-0.97802	-1.38772
H	1.0	-2.92419	-0.82493	1.26334
H	1.0	-0.85917	-2.39098	1.88614
H	1.0	0.04892	-3.54527	-0.36639

Point **n**:

C	6.0	-1.70879	2.59619	0.11445
C	6.0	-1.01285	2.49183	1.47282
H	1.0	-1.16473	3.25563	-0.55933
H	1.0	-2.71463	2.99676	0.23963
H	1.0	-0.96288	3.47426	1.95078
H	1.0	-1.56769	1.84258	2.15321

H	1.0	-1.81066	1.62716	-0.37583
Si	14.0	0.74633	1.86543	1.46229
Cl	17.0	0.87614	-0.17510	1.04517
Cl	17.0	1.58267	2.11923	3.32857
Cl	17.0	1.89585	2.86732	0.07989
Ti	22.0	-0.95934	-0.75729	-0.90350
C	6.0	1.06229	-0.23122	-2.14385
C	6.0	0.55964	-1.48014	-2.55814
C	6.0	-0.78059	-1.27341	-3.03063
C	6.0	-1.10434	0.13224	-2.89495
C	6.0	0.06347	0.75649	-2.32771
C	6.0	-1.75633	-2.80425	-0.00456
C	6.0	-2.57337	-2.25624	-1.04706
C	6.0	-3.10863	-0.98845	-0.59688
C	6.0	-2.58519	-0.78737	0.73282
C	6.0	-1.77949	-1.90277	1.07633
H	1.0	2.03042	-0.06499	-1.69516
H	1.0	1.09743	-2.41546	-2.54159
H	1.0	-1.38193	-2.00926	-3.53701
H	1.0	-1.95175	0.63807	-3.32904
H	1.0	0.18897	1.80966	-2.13189
H	1.0	-1.22124	-3.74094	-0.03479
H	1.0	-2.85378	-2.76749	-1.95241

H	1.0	-3.90722	-0.43156	-1.06013
H	1.0	-2.82995	0.03181	1.39192
H	1.0	-1.22136	-2.01664	1.99383

Point **o**: Same as reaction **II**

CHAPTER 4: MACMOLPLT: A GRAPHICAL USER INTERFACE FOR GAMESS

A paper submitted to the *Journal of Molecular Graphics & Modelling*

Brett M. Bode and Mark S. Gordon

Abstract

A description of MacMolPlt, a graphical user interface for the General Atomic and Molecular Electronic Structure System, GAMESS, is presented. Major features include an input builder for GAMESS; display and animation of molecular structure, normal modes of vibration, reaction paths, orbitals, total electron densities, molecular electrostatic potentials, and density differences. The strategy for direct computation of orbital, total electron density and molecular electrostatic potential surfaces is discussed.

Introduction

In recent years as the speed of computers has increased and the computational methods improved, the complexity of molecular systems studied has steadily increased. Hence, the need for better tools to analyze the results produced by general electronic structure programs such as GAMESS[1] or Gaussian 94.[2] Indeed several commercial programs, Chem3D,[3] HyperChem,[4] Spartan[5] and XMol[6] for example, now have an interface to Gaussian 94.

Since GAMESS is one of the most widely used *ab initio* quantum chemistry programs, MacMolPlt has been developed to provide a user-friendly graphical interface to view the results from GAMESS calculations. MacMolPlt provides a range of visualization

options that are useful to any user. The simple, intuitive interface is particularly useful for introducing novice users to GAMESS. Thus, MacMolPlt is useful in a variety of classroom settings, from teaching students the basic orbital shapes and orbital interactions in their first chemistry course, to introducing the techniques of quantum chemistry to advanced students without requiring an in depth knowledge of GAMESS. MacMolPlt is also very useful to advanced GAMESS users providing insights into complex problems through the use of animations and the visualization of complex 2D and 3D surfaces such as orbitals, total electron densities, or molecular electrostatic potentials.

The remainder of this article is organized as follows: A description of the general capabilities of MacMolPlt is followed by more in depth discussions of the most significant features: supported file formats, GAMESS input generator, direct computation of surfaces, and output visualization. Special emphasis is placed on features that are unique to MacMolPlt.

Program Overview

The primary goal in the development of MacMolPlt was to provide a program to help analyze GAMESS output that is useful for both novices and advanced users. To aid the novice user MacMolPlt also includes an optional input generator for GAMESS that allows the user to set up most of the common GAMESS input options. The advanced user may create input files by hand, perform the calculation and then use MacMolPlt to analyze the results. Thus, MacMolPlt is capable of analyzing results from calculations which are too complex to be set up with the built-in input generator.

For flexibility, MacMolPlt can visualize information from a variety of file formats for input and output. These include all of the files used by GAMESS as well as several other formats as discussed in the section on input file formats. In addition to files, MacMolPlt also allows the user to copy and paste molecular coordinates in several formats, providing a quick way to visualize molecular structures. Input can contain as little as a single atom in one geometry, to many hundreds of atoms in a series of hundreds of geometries, each potentially containing normal modes of vibration and several forms of molecular orbitals. While the current focus of development has been to provide an interface to GAMESS, there is no reason that the results of other programs, such as Gaussian 94, could not also be visualized once the code to interpret the proper file formats is added.

Memory is allocated dynamically in MacMolPlt. Thus, there are no compile time limits on any significant parameter including the number of atoms, the number of geometries, and the number of basis functions. This flexibility is made possible by an object oriented code design. The design allows each geometry point, or frame, to be independent of the other frames. Thus, frames can be easily added or removed and each frame may contain different types and amounts of data. For instance, only the first and last frames in an optimization normally contain molecular orbital vectors (MOs) and only the last frame will contain localized orbitals.

The most powerful features of MacMolPlt come from its output visualization. Not only can the visualization of output be crucial to a correct interpretation of the results, but it can also provide a valuable educational tool. MacMolPlt is tuned to providing real-time visualization with modest hardware requirements. This makes it possible to load onto a standard laptop computer to present results at scientific meetings or in classrooms for

educational demonstrations. In addition to direct onscreen visualization, MacMolPlt can produce output in variety of formats suitable for both hardcopy and web based publication. This is discussed in detail in the section on output file formats.

The output files from GAMESS are read directly and presented for the user to choose which parts to visualize. Visualization options include simple structures, series of structures, normal modes of vibration, molecular orbitals, total electron densities, density differences, and molecular electrostatic potentials. All surface visualization options are implemented directly in MacMolPlt so no further batch processing is required. Since some visualization options require enough CPU time to prevent their being considered truly real time, MacMolPlt allows the user to easily specify the display quality. Thus the user can compute surfaces and images to a low resolution for daily work, and then increase the resolution for publication quality output. This also allows users to tune the graphics level to the performance of their individual computers, from a ten year old computer to today's high end computer with 3D hardware graphics acceleration.

Input File Formats

MacMolPlt supports several text file formats for input. These include the three primary GAMESS output files (log file, dat file and irc file, described below), GAMESS input files, MolPlt mol files, XMol XYZ files, and Protein Data Bank (PDB) files. The last three formats provide compatibility with other programs, but typically include only basic structural information and possibly normal modes. Additional information such as energetics and molecular orbitals currently must come from GAMESS output files. The file parsing code is designed to automatically determine the type of the file selected for reading. Thus the

user is not required to do any hand editing of the input file or perform any intermediate steps in order to read in the results of a GAMESS calculation.

The GAMESS output file containing the most information (the "usual" output for an electronic structure calculation) is the log file. Therefore, it is from the log file that MacMolPlt reads most of the results of the GAMESS calculation: initial geometry, basis set, GAMESS control parameters, molecular wavefunction, normal modes of vibration, localized molecular orbitals, plus all geometries computed as the result of a molecular optimization, intrinsic reaction coordinate (IRC), or dynamic reaction path (DRP) calculation. Results from IRC and DRP calculations, which may take several computational runs to complete, can be easily merged into one file.

In addition to the GAMESS log file MacMolPlt also reads information from the GAMESS irc and dat files. The irc file contains the basic structure and energy information for IRC and DRP calculations. Since these calculations usually result in large numbers of geometries it is advantageous to store only the irc file since it is much smaller than the log file for the same run. The dat file is used by GAMESS to store formatted information that would be useful to restart a calculation, such as molecular orbitals, gradient, and hessian. MacMolPlt uses the dat file as an alternative source for the molecular orbitals.

MacMolPlt also reads in the input files for the MolPlt program, a 2D X-windows program included with GAMESS. These files contain the molecular structure and optionally one or more vibrational normal modes. For compatibility with other programs the XMol style XYZ format can be used. The XYZ format is very compact, including only the cartesian coordinates and possibly a single vibrational normal mode. Multiple geometries may be concatenated together, so it is a useful format for transferring a reaction animation from one

program to another. Finally, files in the Protein Data Bank can also be used for input, though only the basic atomic coordinates are used by MacMolPlt.

Output File Formats

There are two purposes to the output file formats provided by MacMolPlt. The first is to create input files for GAMESS. This feature of MacMolPlt will be discussed in detail in the section on GAMESS Input Generation. The second purpose is to provide high quality graphics for publication either on the web or as hardcopy.

Hardcopy may be produced by two options. First the desired image may be copied and pasted into another application for annotation before printing. This method works well for 2D vector graphics, but can problematic for bitmaps, since many applications cannot handle the very large bitmaps (possibly several megabytes in size) needed to produce acceptable quality when printed. The user may also print directly to any printer, or to a postscript file suitable for later printing on any postscript printer.

Publication on the web has somewhat different requirements than hardcopy output. The graphics resolution on the web is usually limited to 72 dots per inch (dpi), and the size of the images must be kept to a minimum to ensure a reasonable download time. To achieve small file size MacMolPlt uses three different file formats. The first is the XYZ format. As described in the previous section, the XYZ format is quite compact and very useful for transferring one or more structures between programs or for web display. The downside of this format is that the publisher has little control over what the viewer sees since the image is rendered upon download. The viewer must also have a browser plugin installed capable of viewing XYZ files.

To provide images viewable on a wider variety of computers MacMolPlt makes use of the QuickTime software libraries.[7] Using QuickTime, MacMolPlt can produce still images in either the QuickTime movie format or in the jpeg still image format. MacMolPlt also exports animations, including normal mode animations and reaction path animations, into QuickTime movie format. Movies of reaction path animations can include a simple X-Y plot of the energy or other property such as bond lengths or bond angles. The graph gives the viewer a better understanding of how the currently viewed structure relates to the other points in the animation. To help minimize the size of the movie file, MacMolPlt makes use of both spatial (per frame) and temporal (based on a series of frames) compression.

Input Generation

MacMolPlt incorporates a basic input generator to aid users, particularly novice users, to produce input files for GAMESS. However, the use of the input builder is completely optional. MacMolPlt's ability to correctly parse GAMESS output files is not dependent on the use of the input builder. Figure 1 illustrates the basic interface of the input builder. Currently many, but not all GAMESS input groups have been programmed into the input builder. The input groups that are present allow the user to specify the type of wavefunction, basis set, type of run (optimization, IRC, saddle point search, etc), molecular parameters (charge, multiplicity, symmetry, etc), and various other options which affect how GAMESS runs the calculations (amount of memory, maximum CPU time, etc). In addition, the input builder can provide assistance to even advanced users by making it easy to include groups such as optimized MO vectors from a previous calculation, or a list of Z-matrix variables

which tend to be tedious to input by hand. The molecular structure can also be easily modified making it easy for the user to create an isomer or add substituents.

As the user makes choices such as the wavefunction type, MacMolPlt adjusts the available options to be consistent with the capabilities of GAMESS. For example, if the run type is set to optimize and the SCF type to restricted open shell Hartree-Fock (ROHF) the option to use Møller-Plesset second order perturbation theory (MP2) will be disabled since GAMESS cannot currently perform gradients with MP2 wavefunctions. However, if the SCF type is set to RHF and MP2 is also selected then an additional group will appear in the optional groups list for setting MP2 specific options. These optional groups, when selected, provide a pane with a list of options that normally default to an appropriate setting, but still might be of interest. For example, if the user desires to reduce disk space usage they can choose the SCF Options pane and turn on the Direct SCF option.

As a further aid to the new GAMESS user, the input builder also provides online help by means of the balloon help system. When activated via a menu item, helpful messages are displayed next to the item the cursor is currently over. The help message includes a brief description of the option along with the GAMESS group and keyword corresponding to this option. For example the help text for the "Use MP2" option shown in Fig. 1 is: "$CONTRL:MPLEVL - Click to use 2nd order Møller-Plesset perturbation theory. Implemented for RHF energies and gradients and open shell energies." Thus the user can easily associate keywords in the resulting input file with selections in the interface. This also provides information to the user regarding where to find more detailed information in the GAMESS manual.

Surfaces

The four supported surface types in MacMolPlt are Orbitals, Total Electron Densities (TED), Molecular Electrostatic Potentials (MEP), and any general grid of data read in from a file. Each of the supported surface types can be viewed as either a 2D contour map in a user-defined plane, or as a 3D isosurface. Each surface is independent of all other surfaces both within the same frame and for different frames, although it is quite easy to apply the same surface parameters across all frames creating an animation of that surface. In addition a single frame can have multiple surfaces of the same or different types. While any number of surfaces can be visible at the same time, it is unusual for more than two or three surfaces to be visually useful at once. A more useful advantage to having multiple surfaces is to set up surfaces with the optimum settings, then make them invisible. Later, they can be shown quickly one at time for presentation purposes. The surface display is independent of the surface type and is discussed in more detail in the next section. One key feature of MacMolPlt is that all 3 major surface types are computed directly by MacMolPlt on the local CPU(s). This means that the user can visualize any desired surface without waiting on a batch process or other intermediate step.

The orbital surface supports the following types of orbitals: atomic orbitals (AOs), molecular eigenvectors including UHF alpha and beta sets and MCSCF or GVB optimized and natural orbitals, and localized molecular orbitals. Figure 2 illustrates the options available for 2D orbital surfaces. Figure 3 illustrates a corresponding contour map in the plane of the CH_2 molecule. The user is free to choose any orbital from any of the orbital sets that have been read in from GAMESS output. The example in Fig. 2, which is an MCSCF calculation on CH_2, includes the MCSCF natural and optimized orbitals and the atomic

orbitals as specified in the basis set. For MOs the user may list the orbital occupation numbers, as shown in Fig. 2, or the orbital energies. The orbital symmetries are also shown if they are known. One particularly nice feature is that when the user selects a molecular orbital, the MO vector is listed. This allows users to see the orbital visually and observe how the orbital is numerically constructed from the individual atomic basis functions.

The total electron density surface type is available for any calculation that included the natural orbitals (the orbitals for which the density matrix is diagonal) in the output such that the orbital occupations are known. Since the goal is to achieve real time performance, the order of computation is very important. The method illustrated in Figure 4 works quite well. The key to this method is that AO amplitudes are calculated the minimum possible number of times. To accomplish this the computation of the AO amplitudes is done before the loop over MOs and is stored in a one-dimensional array of the same dimension as the basis set. Then, within the loop over the occupied MOs, the MO coefficients are multiplied by the AO amplitudes to produce a MO amplitude. The MO amplitude is squared and multiplied by the MO occupation number to obtain the electron density for that orbital. The individual MO densities are summed to produce the total electron density. In addition to plain TED plots, MacMolPlt can also color map the MEP to the surface of 3D contours. To reduce the time required to compute the MEP the calculation is performed only on the actual surface points. Thus, the MEP calculation is done after the TED grid is computed, contoured and the contour is reduced to the unique 3D points.

MacMolPlt also allows computation of MEP surfaces. A MEP is defined by the equation:

$$V(r) = \sum_{A} \frac{Z_A}{|r - R_A|} - \sum_{\mu,\nu} P_{\mu\nu} \int \frac{\varphi_\mu \varphi_\nu}{|r - r'|} dr'$$

where Z_A is the nuclear charge on atom A centered at R_A, $P_{\mu\nu}$ is the density element corresponding to the basis functions φ_μ and φ_ν. While this integral is not difficult or expensive to compute, it does depend on r and thus must be computed separately at each grid point. This can result in substantial calculation times for large molecules and 3D grids. However, the performance is quite acceptable for 2D grids and will only get better as the speed of CPU's increases.

The fourth surface type supported by MacMolPlt is a simple grid read in from a file. Thus, it can be used to display any arbitrary property. However, because grids can be squared, added, and subtracted as they are read from file, the main use is for density differences. All surface types incorporate the ability to export grids and to synchronize grid parameters. Thus, it is possible to create density differences for individual orbitals (of any orbital type) as well as for the total electron density.

Graphics

The goal of the graphics in MacMolPlt is to provide visual results with sufficient quality to be useful, while keeping the speed high enough to provide smooth real-time model rotation. This goal is accomplished using two separate drawing engines. The first uses simple 2D vector graphics to visualize molecular structures, normal modes, and 2D contour maps. The second uses a true 3D display to render everything available in the 2D mode plus complex 3D surfaces. Real time rotation is provided in either drawing mode via a virtual 3D-trackball scheme[8].

Figure 5 illustrates the basic 2D drawing mode, which utilizes a rendering engine based on 2D vector graphics which are z-buffered based on the atomic coordinates in the screen orientation. This engine provides high-speed graphics on virtually any computer made in the last ten years, while providing sufficient detail for normal modes with ball and stick model display. Also, due to the small size required to save vector-based graphics, it is quite effortless to copy and paste images into other applications for annotation or inclusion into larger documents.

Figure 6 illustrates the true 3D rendering mode, which uses the QuickDraw 3D[9] rendering engine to provide true-3D display with lighting and shading effects. While this graphics engine does require considerably more CPU power than the vector graphics mode, it does run on most computers built in the last four years with no 3D hardware accelerator required. MacMolPlt will also automatically take advantage of any 3D hardware graphics accelerators available on a particular computer to provide enhanced speed and additional effects such as transparent surfaces.

2D contour maps can be viewed using either display mode. The maps consist of a 2D grid spanning a user-defined plane. Specification of the 2D plane can be cumbersome for the user, thus to avoid this pitfall MacMolPlt uses the plane of the screen to define the plotting plane. Once the desired plane has been found, the plane's orientation can be fixed such that it rotates with the molecule. One advantage of this scheme is that with the plane set to the plane of the screen and with the ability of MacMolPlt to compute the surfaces in real time, the user can scan the plane through the molecule, allowing the user to locate areas of special interest very quickly. The plane of the screen may also be easily set to a plane defined by any three

atoms or defined by the cartesian axes. One time and memory saving feature is that the contours are displayed as the grid is contoured. Thus the individual contours are not stored.

A 3D grid spanning a volume of user customizable size and resolution defines all 3D surfaces, viewable only in the true 3D mode of display. Once the grid has been computed or read in from a file, it is contoured into a 3D surface of constant value (an isosurface) using an improved marching cube algorithm.[10] Vectors perpendicular to the surface are computed to smooth out solid surface display. Then the surface is submitted to the 3D-drawing engine for display as either a solid (opaque or transparent) or wire frame surface. The colors of the positive and negative contours of each surface can be set independently such that multiple surfaces can be distinguished from each other if displayed simultaneously. Since the grid is preserved, the user can easily and quickly scan the possible isosurface values to find the value that the user deems best.

Animations

One very important feature of MacMolPlt is its ability to animate computations. This includes the animation of normal modes as well as any multi-frame computation from GAMESS. However, the geometry alone does not provide enough information to truly understand a reaction. Other parameters are also important such as the energy (the total energy or the kinetic energy for DRP's), gradient, or specific bond lengths or angles. To provide information on these parameters MacMolPlt includes a simple graphing feature that can provide an indication of the relative value of the current frame to the other frames in the animation. Because the simple graphing code is not intended for the creation of publication quality output, the graphed values may be exported as tab-delimited text suitable for most

graphing software. Since an animation is difficult to illustrate on paper a sample figure is not presented here, however, sample animations are available on this journal's online web site[11].

The two primary examples of the utility of animation of a multiple geometry calculation are IRC's and DRP's. The IRC path is a minimum energy path connecting a transition state to its nearest minima. Because each point is significant, it can be very important to view the IRC path to gain a better understanding of how a reaction proceeds. MacMolPlt makes viewing the entire reaction path easy by providing built in animation capabilities as well as the ability to splice together multiple computations (e.g. several IRC's for several steps in a complex mechanism) into one smooth animation. Thus an entire reaction sequence can be shown instead of just the stationary points, leading to a better understanding of the entire reaction.

Animations are even more important for viewing the results of a dynamic reaction path calculation, since there are no stationary points computed. By their nature, dynamics calculations require the computation of a very large number of geometries. However, since MacMolPlt imposes no limits on the number of geometries even large runs can be accommodated. Also, since the DRP is often required to compute more geometries for numerical stability than are necessary for a smooth animation, the code can also be configured to skip points as the file is read in from disk. Finally the speed of the animation can be adjusted to suit the number of points in the animation and the user's personal taste.

Conclusions

MacMolPlt provides an easy to use interface for the GAMESS package offering a range of features appealing to the novice as well as the veteran user. While the input is aimed

primarily at new GAMESS users, the powerful output visualization capabilities should appeal to any user. MacMolPlt also bridges the need for a program which is fast enough to use for every computation, yet capable of producing output of sufficient quality for publication. MacMolPlt also supports the most popular formats for publication of results on the web, such as jpeg and quicktime movies. In the future, as the speed of computers and computational methods improves, the need for the powerful visualization techniques provided by programs such as MacMolPlt will only increase as the complexity of molecular systems studied increases.

Availability

MacMolPlt is available for free of charge for all users. To obtain MacMolPlt please refer to the web address: http://www.msg.ameslab.gov/GAMESS/Graphics/MacMolPlt.shtml. Complete system requirements, feature lists, as well as the program itself can be found on that site.

Acknowledgements

The development of MacMolPlt has been supported by grants from the National Science Foundation (CHE-9633480), Department of Defense CHSSI program, Iowa State University in the form of a Department of Education GAANN fellowship awarded to B.M.B., and Apple Computer, Inc. which provided a computer for code development.

References

1. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Jr Montgomery, The general atomic and molecular electronic structure system. *J. Comp. Chem.*, **14**, 1347 (1993).
2. Gaussian 94, M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Replogle, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewert, M. Head-Gordon, C. Gonzalez, and J.A. Pople, Gaussian Inc., Pittsburg PA, 1995.
3. CambridgeSoft Corporation, Cambridge, MA.
4. HyperCube, Inc., Gainsville, FL.
5. Wavefunction, Inc., Irvine, CA.
6. Network Computing Services, Inc., Minneapolis, MN.
7. Apple Computer, Inc., Cupertino, CA, See: http://www.apple.com/quicktime/
8. a) M.Chen *Develop: The Apple Technical Journal*, **14**, 40, (1993); b) M. Chen, S.J. Mountford, A. Sellen *ACM. Siggraph Computer Graphics*, **22**, 121, (1988).
9. Apple Computer, Inc., Cupertino, CA, See: http://www.apple.com/quicktime/qd3d/
10. W.E. Lorensen, H.E. Cline *ACM. Siggraph Computer Graphics*, **21**, 163, (1987).

11. The animations and other supplementary material are available at:

http://www.msg.ameslab.gov/Group/Supplementary_Material/MacMolPlt/

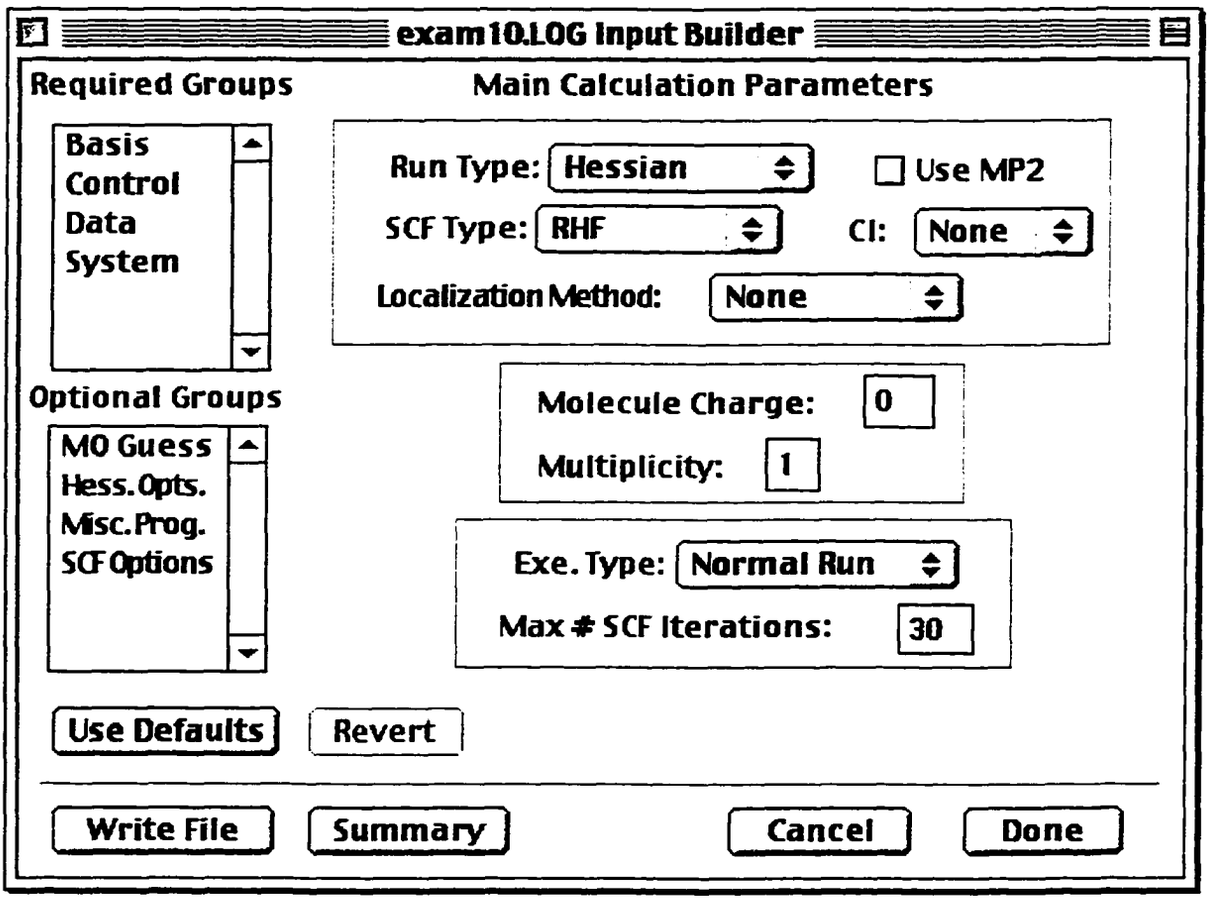


Figure 1: GAMESS input generator

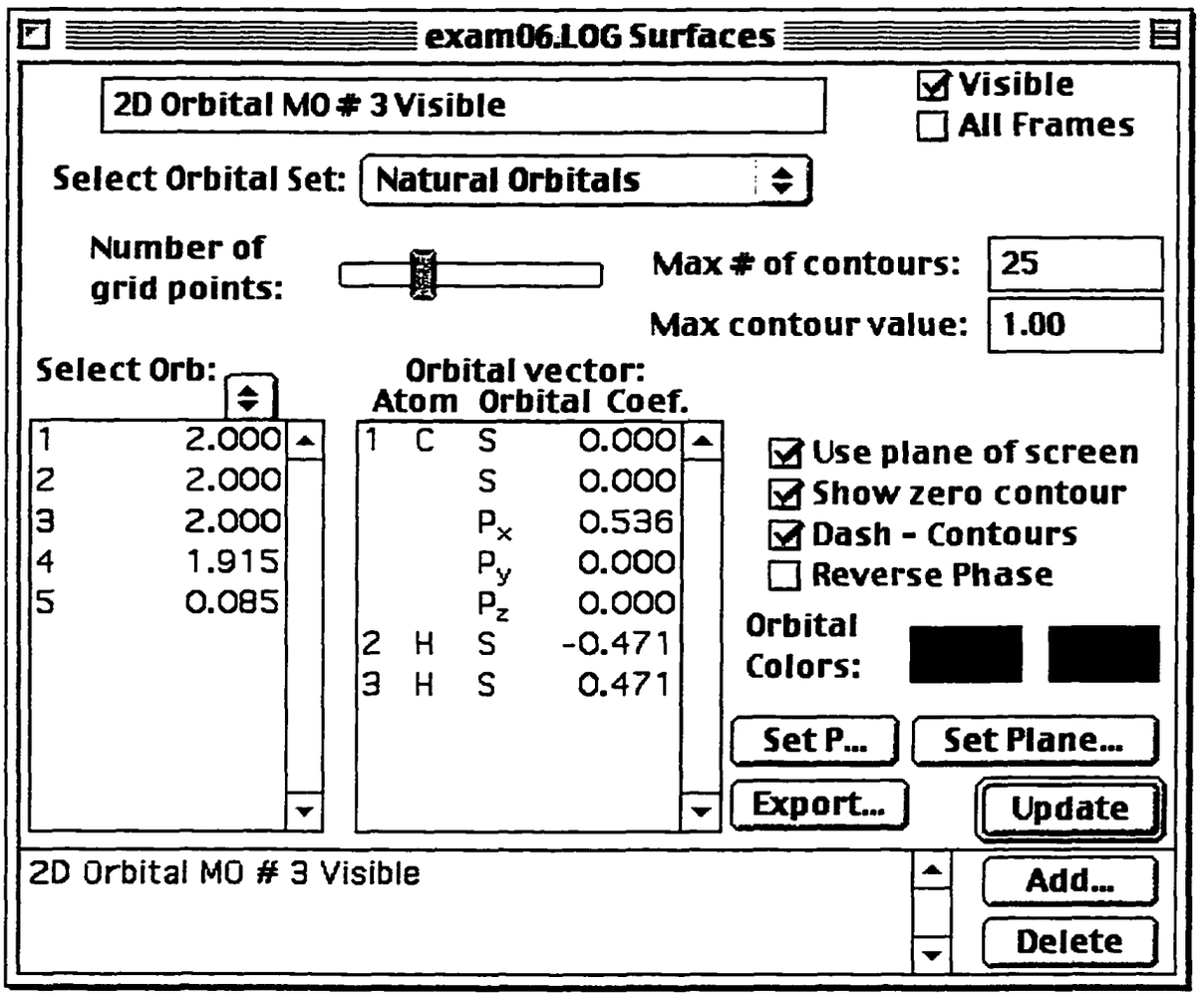


Figure 2: Orbital surface options

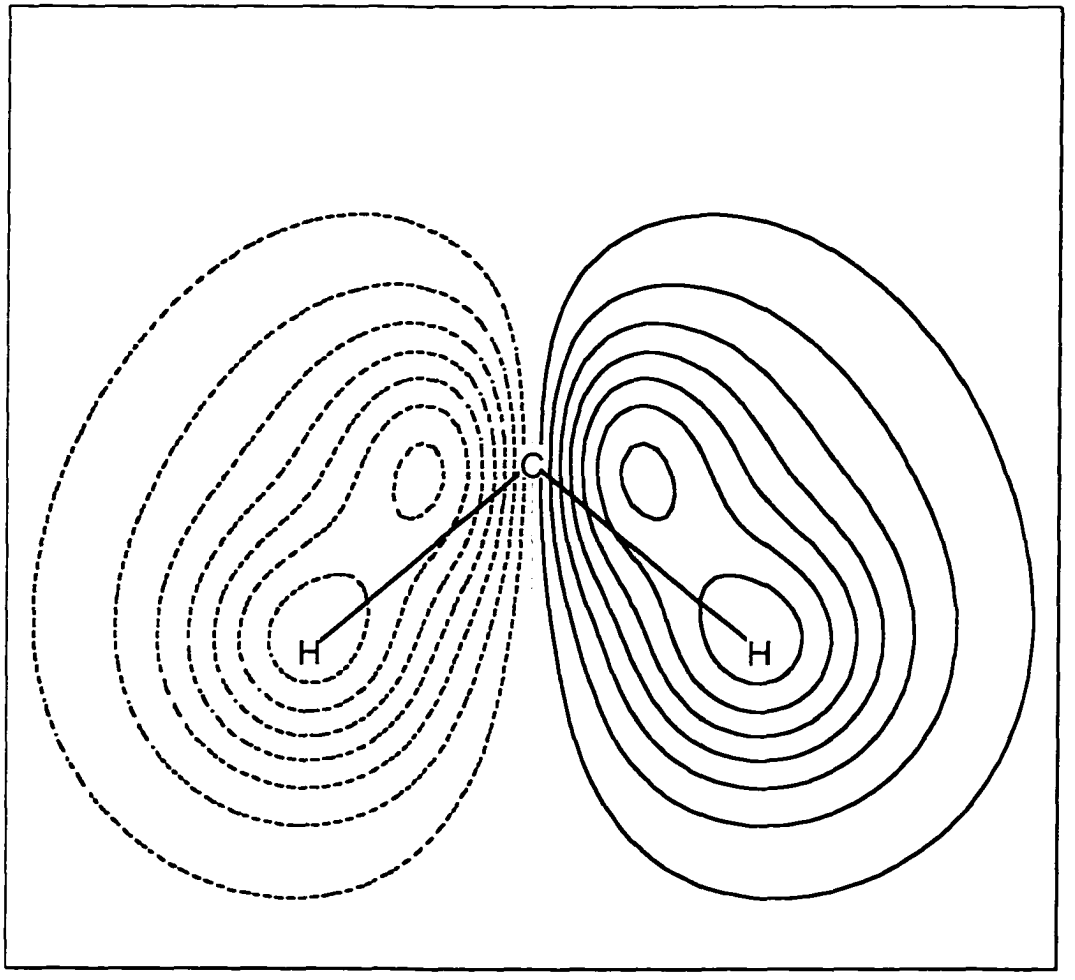


Figure 3: 2D Orbital contour map for CH_2

```
Loop over grid points {
      Loop over basis functions (or AOs) {
            Compute and store AO amplitudes
      }
      DensitySum = 0.0;
      Loop over occupied MOs {
            MODensity = 0.0;
            Loop over basis functions {
                  MODensity += MOCoefficient[I] * AOAmplitude[I];
            }
            DensitySum += MODensity * MODensity * MOOccupation;
      }
      Grid[n] = DensitySum;
}
```

Figure 4: Pseudocode for fast generation of a TED grid

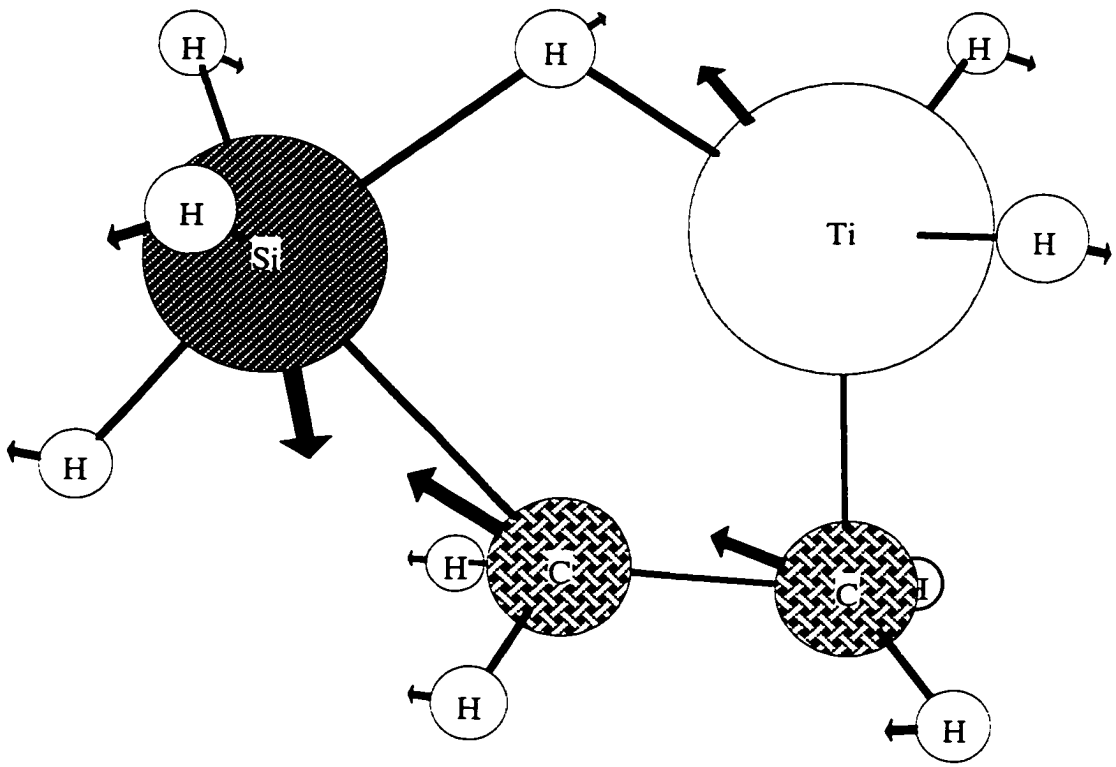


Figure 5: Basic 2D drawing illustrating the molecular structure of a transition state with its corresponding normal mode

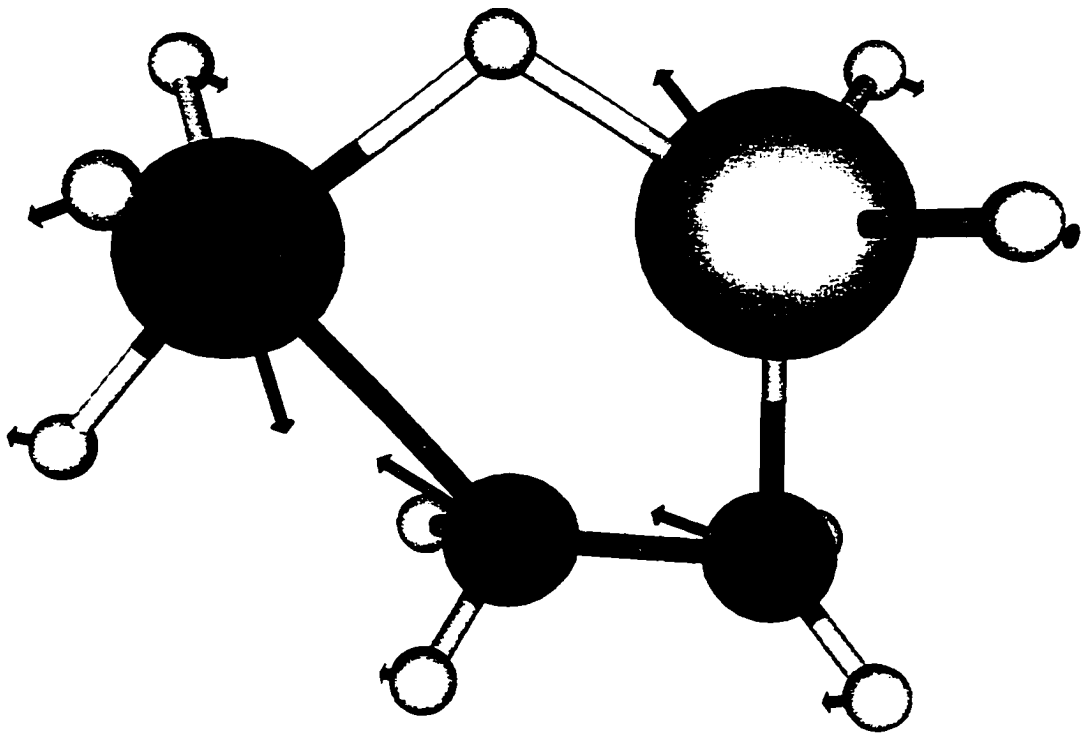


Figure 6: True 3D display utilizing the QuickDraw 3D graphics library

CHAPTER 5: FAST COMPUTATION OF ANALYTICAL SECOND DERIVATIVES WITH EFFECTIVE CORE POTENTIALS: APPLICATION TO SI_8C_{12}, GE_8C_{12} AND SN_8C_{12}

A paper prepared for submission to the *Journal of Chemical Physics*

Brett M. Bode and Mark S. Gordon

Abstract

An improved method is described for the computation of integrals involving effective core potentials. The improved method provides better scalability to higher angular momenta as well as improved speed. The new method is also applied to the determination of the minimum energy structures of Si_8C_{12}, Ge_8C_{12} and Sn_8C_{12}, main group analogs of the Ti_8C_{12} compounds (known as metcars). Relative energies, geometries, and vibrational frequencies are reported for several novel structures.

I. Introduction

The use of Effective Core Potentials (ECP)[1] has grown rapidly in recent years as interest in compounds containing elements from the third and subsequent rows of the periodic table has increased. This increase in use has sparked a renewed interest in improving the efficiency of computations involving ECPs. One of the most significant factors influencing the performance of ECPs is the availability of analytic derivatives. Due to the complicated form of the ECP, it is not easy to directly derive analytic derivatives. However, this difficulty was overcome soon after the introduction of the ECP through the application of translational invariance[2,3] to obtain an analytic formula for the energy first derivative

(gradient).[4] Later this derivation was extended to include energy second derivatives (hessians) by Komornicki *et al.*,[5] and has now been implemented by several groups.[6,7] The result of these derivations are the following formulae for the energy, gradient and hessian involving ECPs:

$$\text{Energy } E_{ECP} = \sum_{\mu}\sum_{\nu}\left\langle \phi_{\mu}\left|\sum_{C}\hat{u}_{C}\right|\phi_{\nu}\right\rangle \tag{1}$$

$$\text{Gradient } E^{a}_{ECP} = 2\sum_{\mu}\sum_{\nu}D_{\mu\nu}\left\langle \phi^{\alpha}_{\mu}\left|\delta_{AI}\sum_{C}\hat{u}_{C} - \hat{u}_{A}\right|\phi_{\nu}\right\rangle \tag{2}$$

$$\begin{aligned}\text{Hessian } E^{ab}_{ECP} = 2\sum_{\mu}\sum_{\nu}D_{\mu\nu}\Bigg(&\left\langle \phi^{\alpha\beta}_{\mu}\left|\sum_{C}\hat{u}_{C}\delta_{AI}\delta_{BI} - \hat{u}_{B}\delta_{AI} - \hat{u}_{A}\delta_{BI} + \hat{u}_{A}\delta_{AB}\right|\phi_{\nu}\right\rangle \\ &+ \left\langle \phi^{\alpha}_{\mu}\left|\sum_{C}\hat{u}_{C}\delta_{AI}\delta_{BJ} - \hat{u}_{B}\delta_{AI} - \hat{u}_{A}\delta_{BJ} + \hat{u}_{A}\delta_{AB}\right|\phi^{\beta}_{\nu}\right\rangle\Bigg)\end{aligned} \tag{3}$$

E^{a} refers the derivative of the energy with respect to coordinate a on center A; ϕ_{μ} and ϕ_{ν} are basis functions located on center I and ϕ^{α}_{μ} is the derivative of ϕ_{μ} with respect to a coordinate of center I; $D_{\mu\nu}$ is an element of the density matrix; $\hat{u}_{C}$ is the ECP on center C. Since the derivatives of gaussian type functions are simply linear combinations of gaussian functions with different angular momentum, these equations all represent the same basic ECP integral. However, a derivative of a gaussian function raises the maximum angular momentum by one for first derivatives and two for second derivatives. Thus, in order to compute the second derivative of a g-type basis function, basic integrals must be computed over h and i-type functions. Due to the number of integrals with high angular momenta it is important to compute the ECP intregrals efficiently. Since the angular portions of the ECP integrals have simple analytic solutions, the main difficulty in the ECP calculation is the computation of the radial integrals.

II. Background

There are currently two methods in use for the computation of the radial integrals. The first is the original implementation by Kahn as described in ref. 1. The second method was proposed by McMurchie and Davidson[8] and has since been improved by several groups.[9] However, a third method, proposed by Kolar,[10] when combined with the original method of Kahn provides a very efficient method which easily scales to higher angular momenta. Our modifications to the original method of Kahn involve the solution of the general type 2 radial integrals given by the equation (in the notation of ref. 8):

$$Q^{N}_{\lambda\lambda'}(k_A, k_B, \alpha) = \int_0^\infty dr\, r^N e^{-\alpha r^2} M_\lambda(k_A r) M_{\lambda'}(k_B r) \tag{4}$$

M_λ is a modified spherical bessel function of the first kind; N is defined to be the sum of the powers on X, Y, and Z for both gaussian centers, also referred to as κ, plus the power of r from the ECP projector, n_{kl} (which may range from 0 for s-projectors to 4 for g-projectors). Thus $N=\kappa+n_{kl}$. Due to symmetry, the matching angular integrals vanish unless $\kappa+\lambda+\lambda'$ is even and $|\lambda-\lambda'|\leq\kappa$. Also, a recursion relationship for the radial integrals may be derived directly from a standard relationship of the modified spherical Bessel functions[11]:

$$Q^{n}_{\lambda\lambda'} = Q^{n}_{\lambda-2,\lambda'} - \frac{2\lambda-1}{k_A} Q^{n-1}_{\lambda-1,\lambda'} \tag{5}$$

Using these relationships the number of radial integrals that must be explicitly calculated may be greatly reduced. In fact, the majority of the integrals involve κ=0, and λ=λ'. Thus consider:

$$Q^{n+u}_{\lambda\lambda}(k_A, k_B, \alpha) = \int_0^\infty dr\, r^{n+u} e^{-\alpha r^2} M_\lambda(k_A r) M_\lambda(k_B r) \tag{6}$$

It is now useful to expand the modified spherical Bessel functions using:

$$M_k(x) = (-1)^{k+1} \sum_{i=1}^{k} \frac{\alpha_{ik}}{x^i} \cosh(x) + (-1)^k \sum_{i=1}^{k} \frac{\beta_{ik}}{x^i} \sinh(x) \tag{7}$$

where α and β are integer coefficients defined in ref 11. General Q integrals then reduce to combinations of the following:

$$\rho_n = \int_0^\infty e^{-cr^2} \sinh(k_A r) \sinh(k_B r)\, r^n\, dr \tag{8a}$$

$$\sigma_n = \int_0^\infty e^{-cr^2} \sinh(k_A r) \cosh(k_B r)\, r^n\, dr \tag{8b}$$

$$\overline{\sigma}_n = \int_0^\infty e^{-cr^2} \cosh(k_A r) \sinh(k_B r)\, r^n\, dr \tag{8c}$$

$$\tau_n = \int_0^\infty e^{-cr^2} \cosh(k_A r) \cosh(k_B r)\, r^n\, dr \tag{8d}$$

It is also convenient to make the following definitions:

$$A_n(c, a_i) = \int_{-a_i}^\infty e^{-cx^2} (x + a_i)^n dx \tag{9a}$$

$$a_1 = \frac{k_A + k_B}{2c} \tag{9b}$$

$$a_2 = \frac{|k_A - k_B|}{2c} \tag{9c}$$

$$\Gamma_i = \frac{1}{4} e^{ca_i^2} \tag{9d}$$

$$B_{n,i} = A_n(c, a_i) + A_n(c, -a_i) \tag{10a}$$

$$C_{n,i} = A_n(c, a_i) - A_n(c, -a_i) \tag{10b}$$

The above integrals then become:

$$\rho_n = \Gamma_1 B_{n,1} - \Gamma_2 B_{n,2} \tag{11a}$$

$$\tau_n = \Gamma_1 B_{n,1} + \Gamma_2 B_{n,2} \tag{11b}$$

$$\sigma_n = \Gamma_1 C_{n.1} + \text{sign}(k_A - k_B)\Gamma_2 C_{n.2} \tag{11c}$$

$$\overline{\sigma}_n = \Gamma_1 C_{n.1} - \text{sign}(k_A - k_B)\Gamma_2 C_{n.2} \tag{11d}$$

The solution of these integrals will involve three special functions: the Error function:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{12}$$

the Dawson function:

$$d(x) = e^{-x^2} \int_0^x e^{t^2} dt \tag{13}$$

and the Hybrid Dawson-error function:

$$h(x) = e^{-x^2} \int_0^x e^{t^2} \text{erf}(t)\, dt \tag{14}$$

III. Radial integral derivation

Since the most important integrals have $\kappa=0$ and $\lambda=\lambda'$; let us begin by examining the $\lambda=\lambda'=0$ integral.

$$\begin{aligned} Q^{n_{kl}}_{0.0}(k_A, k_B, c) &= \int_0^\infty dr\, r^{n_{kl}} e^{-cr^2} M_0(k_A r) M_0(k_B r) \\ &= \int_0^\infty dr\, r^{n_{kl}} e^{-cr^2} \left[\frac{\sinh(k_A r)}{k_A r}\right]\left[\frac{\sinh(k_B r)}{k_B r}\right] \\ &= \int_0^\infty dr\, \frac{e^{-cr^2}}{k_A k_B}\, r^{n_{kl}-2} \sinh(k_A r)\sinh(k_B r) \\ &= \frac{1}{k_A k_B}\, \rho_{n_{kl}-2} \end{aligned} \tag{15}$$

Since $n_{kl} \geq 0$ we need ρ_n integrals for $n=(n_{kl}-2) \geq -2$. Sinec $\rho_n = \Gamma_1 B_{n.1} - \Gamma_2 B_{n.2}$ = a function of $B_{n.l.}$ so consider B_n integrals.

$$B_{n,i} = \int_{-a_i}^{\infty} e^{-cx^2} (x + a_i)^n dx + \int_{a_i}^{\infty} e^{-cx^2} (x - a_i)^n dx$$
$$C_{n,i} = \int_{-a_i}^{\infty} e^{-cx^2} (x + a_i)^n dx - \int_{a_i}^{\infty} e^{-cx^2} (x - a_i)^n dx \qquad (16)$$

or the integrals can be combined by making appropriate variable subsitutions ($t=x\pm a_i$) to give:

$$B_{n,i} = \int_{0}^{\infty} e^{-c(t^2+a_i^2)} \left[e^{2cta_i} + e^{-2cta_i} \right] t^n dt$$
$$C_{n,i} = \int_{0}^{\infty} e^{-c(t^2+a_i^2)} \left[e^{2cta_i} - e^{-2cta_i} \right] t^n dt \qquad (17)$$

For n=0 both b and c are easy to solve:

$$B_{0,i} = 2 \int_{0}^{\infty} e^{-cx^2} dx = \sqrt{\frac{\pi}{c}} \qquad (18a)$$

$$C_{0,i} = 2 \int_{0}^{a_i} e^{-cx^2} dx = \sqrt{\frac{\pi}{c}} \, \mathrm{erf}(a_i \sqrt{c}) \qquad (18b)$$

For n > 0, differentiating eqs (16) by a_i gives:

$$\frac{dB_{n,i}}{da_i} = n \, C_{n-1,i} \qquad (19a)$$

$$\frac{dC_{n,i}}{da_i} = n \, B_{n-1,i} \qquad (19b)$$

For all n, differentiating eqs (17) by a_i gives:

$$\frac{dB_{n,i}}{da_i} = 2c\left[C_{n+1,i} - a_i B_{n,i}\right] \tag{20a}$$

$$\frac{dC_{n,i}}{da_i} = 2c\left[B_{n+1,i} - a_i C_{n,i}\right] \tag{20b}$$

Now integration of (16) (using integration by parts) or by integrating (19), using $C_n(0)=0$, gives:

$$B_{n,0} = 2\int_0^\infty t^n e^{-ct^2} dt = \frac{n-1}{2c} B_{n-1,0}:$$

$$B_{1,i} = \frac{1}{c} e^{-ca_i^2} + a_i\sqrt{\frac{\pi}{c}}\, erf(a_i\sqrt{c})$$

$$C_{1,i} = a_i\sqrt{\frac{\pi}{c}}$$

$$B_{2,i} = \frac{a_i^2+1}{2c}\sqrt{\frac{\pi}{c}} = \frac{1}{2c} B_{0,i} + a_i C_{1,i} \tag{21a-e}$$

$$C_{2,i} = \frac{a_i}{c} e^{-c a_i^2} + \sqrt{\frac{\pi}{c}}\frac{a_i^2+1}{2c} erf(a_i\sqrt{c}) = \frac{1}{2c} C_{0,i} + a_i B_{1,i}$$

Then by induction using (19) and (20) for n≥2 or directly from (17) using integration by parts (integrate t^n), we have:

$$B_{n,i} = \frac{n-1}{2c} B_{n-2,i} + a_i C_{n-1,i}$$

$$C_{n,i} = \frac{n-1}{2c} C_{n-2,i} + a_i B_{n-1,i} \tag{22a,b}$$

Now consider n<0:

For n = -1 (21) may be written as:

$$\frac{dB_{-1,i}}{da_i} + 2ca_i B_{-1,i} = 2c\, C_{0,i} \tag{23a}$$

$$\frac{dC_{-1,i}}{da_i} + 2ca_i C_{-1,i} = 2c\, B_{0,i} \tag{23b}$$

First solve the homogeneous equations (RHS = 0) to give:

$$B_{-1,i} = \beta e^{-ca_i^2}$$
$$C_{-1,i} = \gamma e^{-ca_i^2} \tag{24a,b}$$

Now plug these solutions into eq. 23, assuming that β and γ are functions of a, and solve using Eq. (18):

$$\frac{d(\beta e^{-ca_i^2})}{da_i} + 2ca_i\, \beta e^{-ca_i^2} = 2c\sqrt{\frac{\pi}{c}}\, erf(a_i\sqrt{c})$$

$$\frac{d\beta}{da_i} e^{-ca_i^2} - 2ca_i \beta e^{-ca_i^2} + 2ca_i\, \beta e^{-ca_i^2} = 2\sqrt{\pi}\sqrt{c}\, erf(a_i\sqrt{c})$$

$$\frac{d\beta}{da_i} = 2\sqrt{\pi}\sqrt{c}\, e^{ca_i^2} erf(a_i\sqrt{c}) \tag{25}$$

$$\beta = 2\sqrt{\pi} \int \sqrt{c}\, e^{ca_i^2} erf(a_i\sqrt{c})\, da_i$$

$$\therefore B_{-1} = 2\sqrt{\pi}\, e^{-ca_i^2} \int \sqrt{c}\, e^{ca_i^2} erf(a_i\sqrt{c})\, da_i$$

$$= 2\sqrt{\pi}\, h(a_i\sqrt{c})$$

$$\frac{d(\gamma e^{-ca_i^2})}{da_i} + 2ca_i\gamma e^{-ca_i^2} = 2c\sqrt{\frac{\pi}{c}}$$

$$\frac{d\gamma}{da_i}e^{-ca_i^2} - 2ca_i\gamma e^{-ca_i^2} + 2ca_i\gamma e^{-ca_i^2} = 2\sqrt{\pi}\sqrt{c}$$

$$\frac{d\gamma}{da_i} = 2\sqrt{\pi}\sqrt{c}\,e^{ca_i^2} \tag{26}$$

$$\gamma = 2\sqrt{\pi}\int\sqrt{c}\,e^{ca_i^2}\,da_i$$

$$\therefore C_{-1,i} = 2\sqrt{\pi}\,e^{-ca_i^2}\int\sqrt{c}\,e^{ca_i^2}\,da_i$$

$$= 2\sqrt{\pi}\,d(a_i\sqrt{c}\,)$$

For n = -2 there seem to be two ways to proceed. One is to integrate Eq. (17) by parts. The second is to repeat the integration of Eq. (21). First consider integrating Eq. (17) (for n=-2):

$$C_{-2,i} = \int_0^\infty e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right]t^{-2}dt \tag{27}$$

Using integration by parts choose:

$$u = e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right], \qquad dv = t^{-2}dt \tag{28}$$

Which gives:

$$(-t^{-1})e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right]\Big|_0^\infty - \int_0^\infty(-t^{-1})\{-2cte^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right] +$$
$$e^{-ct(t^2+a_i^2)}\left[2ca_i e^{2cta_i} + 2ca_i e^{-2cta_i}\right]\}dt$$

$$(-t^{-1})e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right]\Big|_0^\infty - \int_0^\infty\{-2ce^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right] +$$
$$2ca_i t^{-1}e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} + e^{-2cta_i}\right]\}dt$$

$$(-t^{-1})e^{-ct(t^2+a_i^2)}\left[e^{2cta_i} - e^{-2cta_i}\right]\Big|_0^\infty - 2cC_{0,i} + 2ca_i B_{-1,i} \tag{29}$$

Thus the latter part is solved, but the first part is indeterminate at 0. By applying L'Hopital's rule the first term can be resolved to give:

$$C_{-2,i} = 4ca_i e^{-ca_i^2} + 2ca_i B_{-1,i} - 2cC_{0,i} \tag{30}$$

Now consider the second option of integrating Eq. (20). If Eq. (20a) is integrated the result is again Eq. (30) which provides a check for that result. If Eq. (20b) is integrated a relationship for $B_{-2,i}$ may be obtained as follows:

$$\frac{dC_{-2,i}}{da_i} + 2ca_i C_{-2,i} = 2cB_{-1,i}$$

$$\frac{d}{da_i}\left(4ca_i e^{-ca_i^2} + 2ca_i B_{-1,i} - 2cC_{0,i}\right) + 2ca_i\left(4ca_i e^{-ca_i^2} + 2ca_i B_{-1,i} - 2cC_{0,i}\right) = 2cB_{-1,i}$$

$$4ce^{-ca_i^2} - 8c^2a_i^2 e^{-ca_i^2} + 2cB_{-1,i} + 2ca_i\frac{dB_{-1,i}}{da_i} - 2c\frac{dC_{0,i}}{da_i} + 8c^2a_i^2 e^{-ca_i^2} +$$

$$4c^2a_i^2 B_{-1,i} - 4c^2a_i C_{0,i} = 2cB_{-1,i}$$

$$2ca_i\left(2cC_{0,i} - 2a_i B_{-1,i}\right) + 4c^2a_i^2 B_{-1,i} - 4c^2a_i C_{0,i} = 0$$

Similarly $B_{-2,i} = 2ca_i C_{-1,i} - 2cB_{0,i}$ (31)

This provides all of the integrals required for the $\lambda=\lambda'=0$ case, but lower values for n are required to compute other integrals. Using Eq. (17) (and Eq. (20) as a check) the results for n = -3, -4, etc can be obtained:

$$B_{-3,i} = 2a_i^2c^2e^{-ca_i^2} + ca_i C_{-2,i} - cB_{-1,i} \tag{32a}$$

$$C_{-3,i} = ca_i B_{-2,i} - cC_{-1,i} \tag{32b}$$

$$B_{-4,i} = \tfrac{2}{3}ca_i C_{-3,i} - \tfrac{2}{3}cB_{-2,i} \tag{33a}$$

$$C_{-4,i} = \tfrac{8}{9}c^3a_i^3e^{-ca_i^2} - \tfrac{2}{3}cC_{-2,i} + \tfrac{2}{3}ca_i B_{-3,i} \tag{33b}$$

From these we can begin to see a pattern from which the following relations may be derived:

$$B_{-n,t} = \left(1+(-1)^{n-1}\right)\frac{(2ca_t)^{n-1}}{(n-1)(n-1)!}e^{-ca_t^2} - \frac{2c}{n-1}B_{-(n-2),t} + \frac{2ca_t}{n-1}C_{-(n-1),t} \qquad (34a)$$

$$C_{-n,t} = \left(1-(-1)^{n-1}\right)\frac{(2ca_t)^{n-1}}{(n-1)(n-1)!}e^{-ca_t^2} - \frac{2c}{n-1}C_{-(n-2),t} + \frac{2ca_t}{n-1}B_{-(n-1),t} \qquad (34b)$$

Thus all of the integrals for $n < 0$ are easily related to $n = -1$ and $n = 0$ integrals similarly to the relationship between $n \geq 2$ integrals and $n = 0$ and $n = 1$ results. So all possible integrals can be solved in terms of relatively few basic integrals.

This method for the computation of the type 2 radial ECP integrals has been implemented into the GAMESS[12] program. This improvement along with other code modifications has resulted in a significant speedup for the ECP portion of calculations. Table 1 gives some representative timings for a Ge_8C_{12} T_d geometry which will be discussed in detail later in this paper. The timings illustrate the speedup over the older version as well as the parallel scalability of the ECP code. Specifically the speedups are about 5.5 for energies. 6.0 for gradients, and close to a factor of 2 for parallel runs on two nodes. In general the speedup is greater for higher angular momentum functions, thus gradients and hessians have a higher speedup than energies. Analytic derivatives of ECP's have also been implemented through g-functions

IV. Characterization of the minimum energy structures of Si_8C_{12}, Ge_8C_{12}, and Sn_8C_{12}

As interest in fullerenes and fullerene derivatives has grown in recent years, interest in the smallest fullerene like carbon cage compound, C_{20}, has also grown. Several recent studies

have examined the minimum energy geometry of C_{20}.[13,14] Since many of the interesting properties of fullerenes come from the addition or substitution of metal atoms it was of great interest when a class of stable 20-atom molecular clusters (8 metal atoms and 12 carbon atoms referred to as metallocarbohedrenes or metcars) were reported. In particular, Guo and coworkers have reported the formation of M_8C_{12}, where M=Ti, V, Zr, Hf, Mo, or W, through the reaction of laser vaporized titanium with a variety of hydrocarbon gases.[15] This discovery has prompted several theoretical studies attempting to predict the structures and energetics for a variety of transition metal metcars.[16] Most of these studies have considered two types of arrangements for the minimum energy structure. The first is a distorted dodecahedral structure in T_h symmetry and the second is a capped tetrahedron in T_d symmetry. Theoretical studies on Ti_8C_{12} have indicated that the T_d geometry is much more stable.[16]

Considering the similar electronic structure of main group elements such as silicon, germanium, and tin and transition metals titanium, zirconium, and hafnium (s^2p^2 versus s^2d^2) it is reasonable to expect similar compounds to exist. Two studies have considered Si_8C_{12};[17] however, in both of these the only geometry considered was T_h. Thus this work will examine silicon, germanium and tin metcars, with emphasis on the minimum energy structure and lowest electronic states for each system.

V. Computational Methods

All structures were initially optimized using the restricted Hartree-Fock (RHF) level of theory for closed shell states and the restricted open shell Hartree-Fock (ROHF) level of theory for open shell electronic states. Since many of the low lying electronic states have

many unpaired electrons in nearly degenerate orbitals, it is likely that a multiconfigurational treatment is needed. Thus, all structures were then fully reoptimized using multiconfigurational self-consistent field (MCSCF)[18] wavefunctions. The MCSCF active space used for all systems was an (8,8) active space. This active space distributes 8 electrons among 8 orbitals with appropriate consideration of spin and symmetry. This choice allows a full description of all open shell orbitals for all spin states through the nonet (8 unpaired electrons) spin state. Due to the substantial difference in bonding between the different isomers it is not possible to make the active space contain exactly the same orbitals. However, the active space was the same size for all systems and therefore the results are directly comparable. The specific active spaces used contained 8 non-bonding orbitals for the T_h isomer, 6 non-bonding plus 1 E-C bond/antibond pair for the D_2 isomer, and 6 non-bonding plus 2 E-C-E three center bonding orbitals for the D_{2h} isomer.

The basis set used was the SBKJC ECP basis set[19] on Si, Ge, Sn, and a 6-31G(d) basis set on C.[20] One set of d-type polarization functions was added to each heavy atom.[21] All structures were optimized using analytic gradients and then confirmed by computing the matrix of energy second derivatives, or hessian, to obtain the harmonic normal modes and corresponding frequencies (each minimum has zero imaginary modes). The calculated frequencies were also used to obtain the harmonic zero-point energies used to convert energy differences to 0 K enthalpy differences.

The GAMESS[12] program was used for all calculations.

VI. Results

In total 4 distinct geometries were considered. In addition to the expected T_h and T_d geometries two other geometries were also located, one with D_2 symmetry and one with D_{2h} symmetry. In each case, the geometry was determined only for the lowest energy spin state for that symmetry. All four geometries are illustrated in Figure 1 with numbers indicating the symmetry unique atoms. The unique bond lengths are listed in Table 2 for each molecule at both RHF/ROHF and MCSCF levels of theory. Note that metal-metal interatomic distances are given for reference even though there are no metal-metal distances short enough to be considered a bond. Table 3 list the relative energies of each species at the RHF/ROHF and MCSCF levels respectively, including the RHF/ROHF zero point energy correction. The corresponding spin multiplicity is also listed. The absolute energies and zero point energy corrections as well as the complete set of cartesian coordinates are also included as an appendix to this chapter.

From Table 3 we conclude that although there are some significant differences in the relative energies between RHF/ROHF and MCSCF levels of theory, the qualitative energy order is the same. The same is also true of the geometries listed in Table 2, which show only minor differences between the RHF/ROHF bond lengths and those optimized using MCSCF wavefunctions. For the T_h and D_2 isomers which were first optimized using high-spin ROHF wavefunctions, the MCSCF wavefunction predicts that the lowest energy state is a low-spin open shell configuration with a similar number of unpaired electrons as the high-spin state. For comparison with the ROHF result the high-spin quintet state was also optimized using an MCSCF wavefunction and is listed in Table 3. The nonet state for the T_h geometry using an (8,8) MCSCF wavefunction is identical to the ROHF reference and is therefore not

repeated in the MCSCF part of the table.

In addition to the expected T_h and T_d structures, Figs. 1a and 1b, a D_2 structure, Fig. 1c, similar to the T_h structure with two of the C_2 units twisted to a linear arrangement between two E atoms, and a D_{2h} structure, Fig. 1d, with all C_2 units arranged in a linear arrangement between pairs of E atoms, were also found. Perhaps surprisingly, the D_{2h} structure is the preferred geometry for the germanium and tin systems, while the T_h geometry is preferred by the silicon system. The D_2 geometry is 35-60 kcal/mol higher than the minimum energy geometry for each system. The T_d geometry is found to be quite high in energy for all three systems and hessians reveled that it is a high order saddle point with several normal modes which break the T_d symmtery. As the metal is changed from Si to Ge and Sn, the T_h geometry becomes much less favorable, while the D_{2h} geometry becomes more favorable.

Table 2 illustrates that as the metal is changed from Si to Ge and Sn the metal-metal (E-E) and metal-cabon (E-C) distances increase, as might be expected from the increasing atomic radius of the metal atom. Perhaps less obvious is the fact that as the radius of the metal atom increases the C-C bond distance stays the same or more often decreases. There is also a dramatic difference in the C-C bond distance between the four isomers. In the T_h isomer the C-C bond distance of approximately 1.34 Å is characteristic of a C-C double bond. However, the C-C distance of 1.21 Å in the D_{2h} isomer indicates a C-C triple bond. Thus, there is a significant difference in the bond character between the different isomers. These compare to a C-C bond distance between 1.4-1.5 Å for C_{20}[13] and a value of 1.4Å for the Ti_8C_{12} metcars.[17]

The D_2 and D_{2h} structures have fewer E-C bonds than the T_h structure. In the T_h

structure every metal atom is directly bonded to 3 carbon atoms. In the D_2 structure the affect of twisting 2 of the C_2 units is to reduce the E-C bond count on 4 of the metal atoms. This reduction results in an increased bond order for the affected C_2 units, as well as a small reduction in the bond length of the remaining E-C bonds. This effect is even greater in the D_{2h} structure where 4 of the metal atoms are directly bonded to 2 carbon atoms. The remaining 4 metal atoms are directly bonded to only 1 carbon atom and bridge a C_2 unit. Thus, the reduction in the number of direct metal-carbon bonds in the D_{2h} structure results in an increased bond order for the C-C bonds and well as a small increase in the bond order for the remaining metal-carbon bonds.

The vibrational frequencies with significant infrared intensities are listed in Table 4. A simple description is also listed for each frequency, but since the majority of the frequencies involve the entire cluster it is necessary to view animations of each vibration to fully understand the motion of the molecule. Except for two systems where a C_2 stretch is infrared active, all of the listed frequencies involve at least four atoms, such as E-C-C-E stretches, and many involve the entire cage. Most of the low frequency vibrations involve the motion of the C_2 as a unit, either in a twisting motion about the center of the C-C bond or as concerted motion in one direction.

VII. Conclusions

In this paper we have presented an improved method for the computation of integrals involving effective core potentials. This improved method has been implemented into the GAMESS program and has been shown to significantly reduce the computational cost of ECP integrals. The improved program has been applied to the determination of the minimum

energy structures of silicon, germanium, and tin analogs of the Ti_8C_{12} metcars. In addition to the two expected structures, two novel previously unreported arrangements have also been predicted.

Acknowledgements

The work described in this paper was supported by grants from the Air Force Office of Scientific Research (F49-620-95-1-0073), the Department of Defense CHSSI program, and Iowa State University in the form of a Department of Education GAANN fellowship awarded to B.M.B.

References

1. L. R. Kahn, P. Baybutt, and D. G. Truhlar, J. Chem. Phys. **65**, 3826, (1976).
2. A. Komornicki, K. Ishida, K. Morokuma, R. Ditchfield, and M. Conrad, Chem. Phys. Lett. **45**, 595, (1977).
3. L. R. Kahn, J. Chem. Phys. **75**, 3962, (1981).
4. K. Kitaura, S. Obara, and K. Morokuma, Chem. Phys. Lett. **77**, 452, (1981).
5. J. Breidung, W. Thiel, A. Komornicki Chem. Phys. Lett. **153**, 76 (1988).
6. T. V. Russo, R. L. Martin, P. J. Hay, A. K. Rappé J. Chem. Phys. **102**, 9315 (1995).
7. Q. Cui, D. G. Musaev, M. Svensson, and K. Morokuma J. Phys. Chem. **100**, 10936 (1996).
8. L. E. McMurchie, and E. R. Davidson, J. Comp. Phys. **44**, 289, (1981).
9. R. M. Pitzer, and N. W. Winter, Int. J. Quantum Chem. **40**, 773, (1991).

10. M. Kolar, Computer Phys. Comm., **23**, 275, (1981).

11. M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions," Dover, New York, 1964.

12. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery Jr., The general atomic and molecular electronic structure system. J. Comp. Chem. **14**, 1347, (1993).

13. Z. Wang, P. Day, R. Pachter Chem. Phys. Lett. **248**, 121, (1996).

14. J. M. L. Martin, J. El-Yazal, J. P. François, Chem. Phys. Lett. **248**, 345, (1996).

15. (a) B.C. Guo, K. P. Kerns, A. W. Castleman, Jr., Science, **255**, 1411, (1992); (b) B. C. Guo, S. Wei, J. Purnell, S. Buzza, A. W. Castleman, Jr., Science, **256**, 515, (1992); (c) S. Wei, B. C. Guo, J. Purnell, S. Buzza, A. W. Castleman, Jr., J. Phys. Chem. **96**, 4166, (1992); (d) B. C. Guo, K. P. Kerns, A. W. Castleman, Jr., J. Am. Chem. Soc. **115**, 7415, (1993); (e) S. F. Cartier, Z. Y. Chen, G. J. Walder, C. R. Sleppy, A. W. Castleman, Jr., Science, **260**, 195, (1993).

16. (a) P. J. Hay, J. Phys. Chem. **97**, 3081, (1993); (b) Z. Lin, M. B. Hall, J. Am. Chem. Soc. **115**, 11165, (1993); (c) I. Dance, J. Am. Chem. Soc. **118**, 6309, (1996).

17. (a) R. W. Grimes, J. D. Gale, J. Chem. Soc. Chem. Commun., 1222, (1992); (b) R.W. Grimes, J. D. Gale, J. Phys. Chem., **97**, 4616, (1993).

18. (a) K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, S. T. Elbert, Chem. Phys. **71**, 41, (1982); (b) B. Lam, M. W. Schmidt, K. Ruedenberg J. Chem. Phys. **89**, 2221, (1985).

19. (a) For Si: Stevens, W. J.; Basch, H.; Krauss, M. J. Chem. Phys. **81**, 6026, (1984). (b) For Ge, Sn: W. J. Stevens, H. Basch, P. Jasien, Can. J. Chem. **70**, 612, (1992).

20. (a) For H: R. Ditchfield, W. J. Hehre, J. A. Pople, J. Chem. Phys., **54**, 724, (1971); (b) For C: W. J. Hehre, R. Ditchfield, J. A. Pople, J. Chem. Phys., **56**, 2257, (1972); (c) The polarization exponents, H $\zeta_p = 1.1$ and C $\zeta_d = 0.8$, are given by: P. C. Hariharan, J. A. Pople, Theoret. Chim. Acta., **28**, 213, (1973).

21. The exponents used were: Si $\zeta_d = 0.364$, Ge $\zeta_d = 0.254$, Sn $\zeta_d = 0.177$

Table 1: ECP timings (in CPU seconds on an IBM RS/6000 m370), other one and two electron integral timings included for comparison purposes

step	old code	new code	new code on 2 nodes
Energy 1e-ints.	6.0	6.0	4.0
Energy ECP ints.	218.0	39.0	21.9
Gradient 1e- ints.	45.6	45.6	22.3
Gradient ECP ints.	1178.8	196.0	90.2
Gradient 2e- ints.	2523.4	2523.0	1253.8

Table 2: Optimized bond lengths for Hartree-Fock (HF) and MCSCF wavefucntions

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}	
Bond	HF	MCSCF	HF	MCSCF	HF	MCSCF
T_h Geometry						
E-E	3.085	3.088	3.193	3.196	3.698	3.645
E-C	1.911	1.913	1.986	1.986	2.334	2.302
C-C	1.351	1.350	1.339	1.338	1.273	1.283
T_d Geometry						
E_1-E_2	3.345		3.503		3.780	
E_1-C	1.967		2.075		2.269	
E_2-C	2.402		2.514		2.695	
C-C	1.256		1.253		1.252	
D_2 Geometry						
E_1-E_2	3.589	3.606	3.676	3.717	3.930	3.926
E_1-E_2'	3.052	3.052	3.186	3.196	3.670	3.700
E_1-C_1	1.877	1.879	1.963	1.976	2.291	2.280
E_1-C_2	1.931	1.940	2.015	2.047	2.359	2.372
E_1-C_3	1.892	1.879	1.969	1.975	2.304	2.294
E_2-C_1	2.935	2.945	2.942	2.929	2.786	2.775
E_2-C_2	1.954	1.951	2.044	2.031	2.272	2.281
E_2-C_3	1.945	1.945	2.040	2.055	2.278	2.274
C_1-C_1	1.215	1.214	1.217	1.218	1.231	1.231
C_2-C_2	1.343	1.340	1.329	1.324	1.276	1.272
C_3-C_3	1.352	1.353	1.337	1.322	1.279	1.277

Table 2 Continued.

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}	
Bond	HF	MCSCF	HF	MCSCF	HF	MCSCF
D_{2h} Geometry						
E_1-E_2	3.217	3.217	3.372	3.394	3.661	3.691
E_1-C_1	1.866	1.866	1.956	1.962	2.165	2.175
E_1-C_2	1.877	1.877	1.954	1.954	2.134	2.134
E_2-C_1	2.116	2.116	2.250	2.280	2.449	2.487
E_2-C_3	1.865	1.865	1.942	1.942	2.128	2.128
C_1-C_1	1.283	1.283	1.272	1.261	1.264	1.253
C_2-C_2	1.208	1.208	1.208	1.208	1.211	1.211
C_3-C_3	1.206	1.207	1.207	1.207	1.211	1.211

Table 3: Relative energies (kcal/mol)

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}	
Geometry	Energy	State	Energy	State	Energy	State
RHF or ROHF Energies						
a T_h	0.0	^9A_g	75.5	^9A_g	161.9	^1A_g
b T_d (a)	259.8	$^1A_1$	172.7	$^1A_1$	123.1	$^1A_1$
c D_2	49.4	^5A	57.7	^5A	35.9	^5A
d D_{2h}	47.1	^1A_g	0.0	^1A_g	0.0	^1A_g
MCSCF Energies						
a T_h	0.0	^1A_g	88.5	^1A_g	165.0	^1A_g
b T_d (a)						
c D_2	39.6	^1A	59.9	^1A	51.6	^1A
c D_2	50.1	^5A	74.7	^5A	54.7	$^5B_2$
d D_{2h}	46.5	^1A_g	0.0	^1A_g	0.0	^1A_g

(a). Saddle point at the RHF level of theory

Table 4: Vibrational Frequencies with an infrared intensity greater than 0.1 Debye^2/(amu Å^2)

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}		
Geometry	Frequency	Intensity	Frequency	Intensity	Frequency	Intensity	type
a T_h	499	1.4	484	4.3	403	3.3	cs
	499	1.4	484	4.3	403	3.3	cs
	499	1.4	484	4.3	403	3.3	cs
	568	4.1	649	0.2	567	0.4	cr
	568	4.1	649	0.2	567	0.4	cr
	568	4.1	649	0.2	567	0.4	cr
	808	3.2	690	0.8			ca
	808	3.2	690	0.8			ca
	808	3.2	690	0.8			ca
					1587	0.2	c2
					1587	0.2	c2
					1587	0.2	c2
c D_2	160	0.5	135	0.2	103	0.1	cb
	166	0.5	144	0.2			cb
	279	0.5	205	0.8	138	0.3	cb
	282	0.1	216	0.5	187	0.3	cs
	317	0.2	216	0.1	195	0.8	cb
	330	0.2			247	0.8	cs
	356	0.2			277	1.0	cr
	360	0.2			278	0.6	cs
	374	0.2	373	0.6	280	0.3	cr
	424	0.6	391	0.4	299	0.7	cs
	475	2.0	406	2.1	332	0.7	cs
	497	1.3	408	0.2			cr

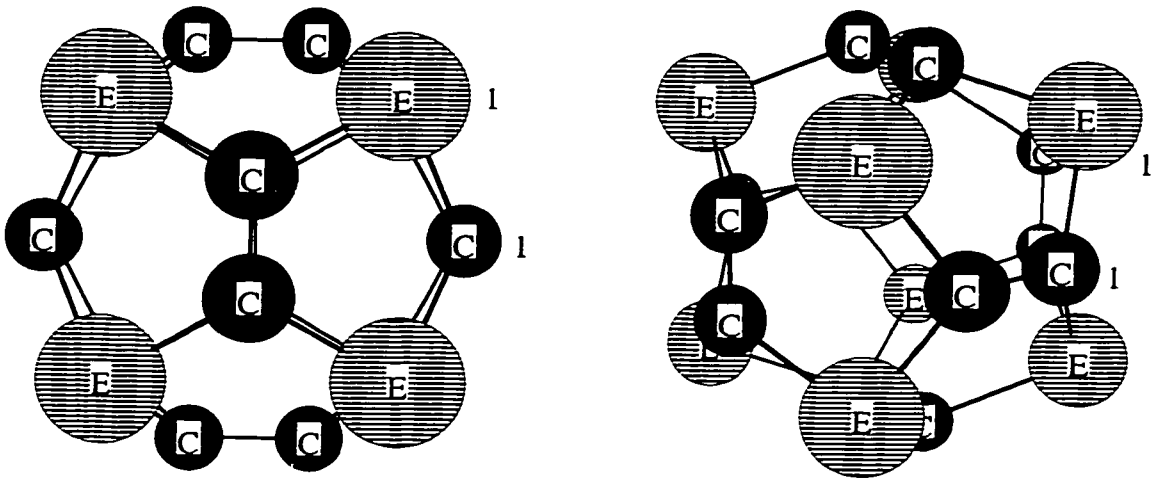
Table 4: Continued.

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}		
Geometry	Frequency	Intensity	Frequency	Intensity	Frequency	Intensity	type
c D_2 cont.	528	1.5	428	0.8	337	1.0	cr
	553	5.0	453	2.1	358	1.4	cs
			457	0.2			cs
			458	1.1	391	4.6	cs
	626	0.2	546	0.4	416	0.6	cr
	641	1.2	558	2.4	450	0.2	cr
	678	2.3	576	4.0	452	2.0	cs
	687	0.6	594	1.1			cs
			604	0.9			cr
			619	0.7			cr
	699	0.2					cs
	715	4.0	686	1.3	454	1.3	cs
	735	0.7					ca
	739	2.5	689	1.9	471	0.6	ca
	805	0.5			479	7.9	cs
	822	5.8	737	0.8	492	0.6	cs
	2295	0.1	1620	0.2			c2
d D_{2h}	262	0.1	233	0.4	188	0.8	eb
	279	0.5	240	0.4	199	0.6	eb
	307	0.7	241	0.2	194	0.3	er
			246	0.2			es
			269	0.2			es
	440	0.6	382	0.3	305	0.3	cr
	538	4.5	464	3.3	370	3.4	cr
	571	0.6	484	12.7	423	8.5	es

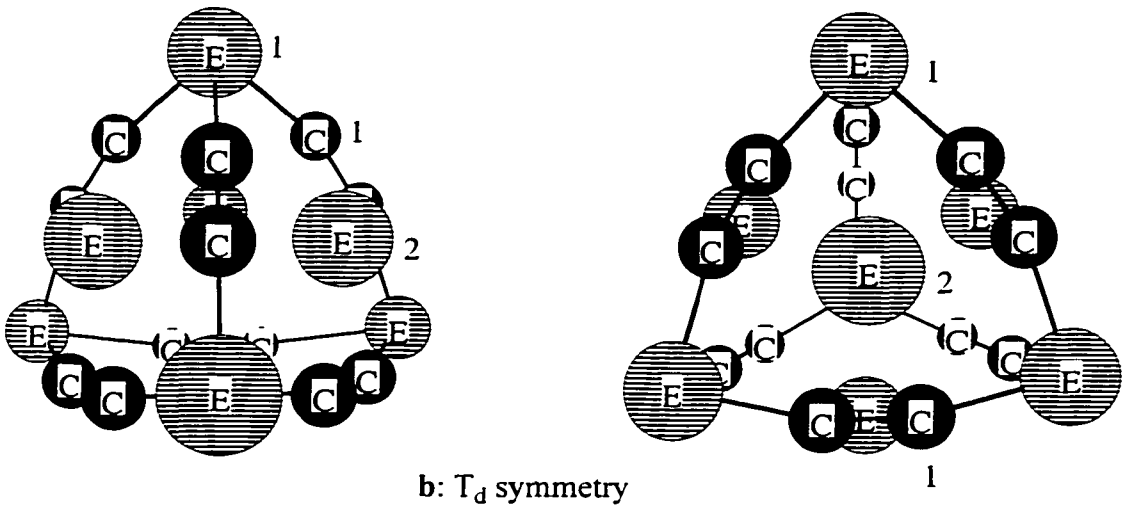
Table4: Continued.

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}		
Geometry	Frequency	Intensity	Frequency	Intensity	Frequency	Intensity	type
d D_{2h} cont.	606	21.3	515	0.7	440	0.6	cr
	723	0.7	600	0.2			es
	746	24.2	612	19.6	526	14.7	es
	780	32.2	653	29.0	584	30.7	es

Codes: cs = symmetric cage stretch, ca = asymmetric cage stretch, cb = cage bend, cr = C_2 rocking motion, c2 = C-C bond stretch, eb = E-C-C-E bend, er = E-C-C-E rock, es = E-C-C-E stretch

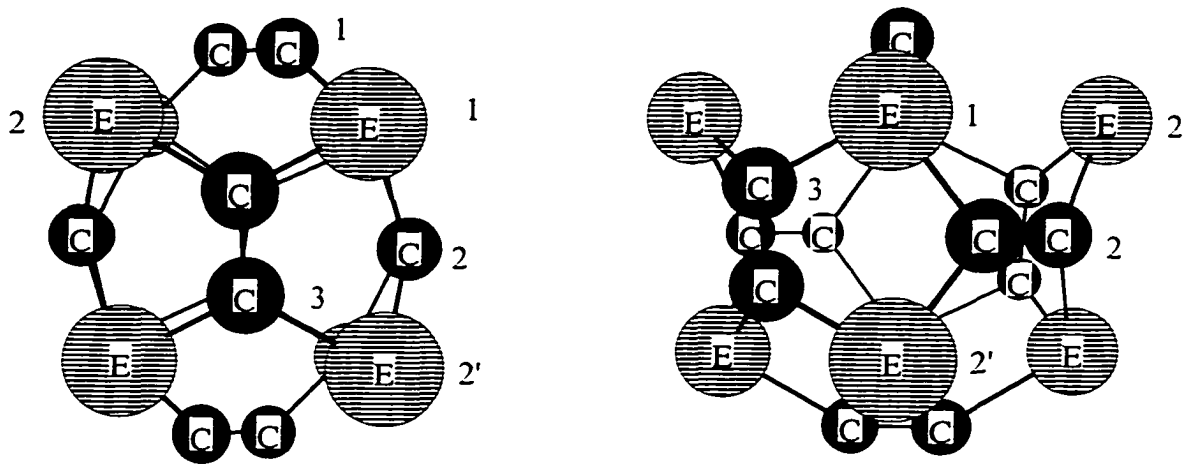


a: T_h symmetry

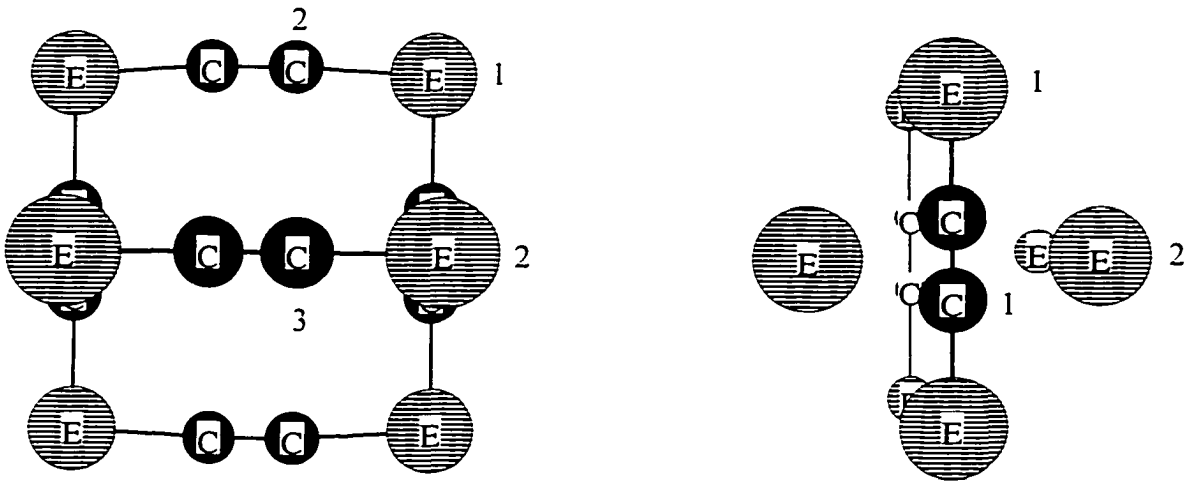


b: T_d symmetry

Figure 1: Structures for E_8C_{12} (where E=Si, Ge, Sn)



c: D_2 Symmetry



d: D_{2h} symmetry

Figure 1: Continued

Appendix

RHF/ROHF absolute energies and zero point energy corrections (in Hartrees)

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}	
Geometry	Energy	State	Energy	State	Energy	State
a T_h	-483.949504	^9A_g	-483.574953	^9A_g	-480.124207	^1A_g
b T_d	-483.509237	$^1A_1$	-483.397641	$^1A_1$	-480.171753	$^1A_1$
c D_2	-483.862607	^5A	-483.596760	^5A	-480.316530	^5A
d D_{2h}	-483.858328	^1A_g	-483.682837	^1A_g	-480.374306	^1A_g
RHF/ROHF Zero Point Energy						
a T_h	0.085117		0.073645		0.063994	
b T_d	0.058792		0.051104		0.049686	
c D_2	0.076872		0.066965		0.055471	
d D_{2h}	0.069066		0.061137		0.056074	

MCSCF absolute energies

	Si_8C_{12}		Ge_8C_{12}		Sn_8C_{12}	
Geometry	Energy	State	Energy	State	Energy	State
a T_h	-483.957271	^1A_g	-483.585616	^1A_g	-480.153802	^1A_g
b T_d						
c D_2	-483.885886	^1A	-483.624615	^1A	-480.325965	^1A
c D_2	-483.869177	^5A	-483.601008	^5A	-480.321024	$^5B_2$
d D_{2h}	-483.867054	^1A_g	-483.714186	^1A_g	-480.408841	^1A_g

Optimized Geometries:

RHF/ROHF Geometries (symmetry unique atoms only):

a T_h

Si_8C_{12}

SILICON 14.0 1.5423678102 1.5423678102 1.5423678102

CARBON 6.0 .0000000000 .6754024883 2.2631670295

Ge_8C_{12}

GERMANIUM 32.0 1.5964372281 1.5964372281 1.5964372281

CARBON 6.0 .0000000000 .6697441419 2.3294671655

Sn_8C_{12}

TIN 50.0 1.7443320473 1.7443320473 1.7443320473

CARBON 6.0 .0000000000 .6682609544 2.5156684452

b T_d

Si_8C_{12}

SILICON 14.0 1.5463130830 1.5463130830 1.5463130830

SILICON 14.0 1.7823778604 1.7823778604 -1.7823778604

CARBON 6.0 0.4439701726 0.4439701726 -2.3171931689

Ge_8C_{12}

GERMANIUM 32.0 1.6318367153 1.6318367153 1.6318367153

GERMANIUM 32.0 1.8569797895 1.8569797895 -1.8569797895

CARBON 6.0 .4428341687 .4428341687 -2.4086432008

Sn_8C_{12}

TIN 50.0 1.7676276160 1.7676276160 1.7676276160

TIN 50.0 1.9980641603 1.9980641603 -1.9980641603

CARBON 6.0 .4426889385 .4426889385 -2.5554066579

c D_2

Si_8C_{12}

SILICON 14.0 1.4413793331 -1.5744365460 -1.6382269639

SILICON 14.0 -1.8087435124 -1.4418081497 -1.9296506844

CARBON 6.0 -.0690704932 -.6721928278 -2.3333232192

CARBON 6.0 2.0760046002 .0279466508 .6708359441

CARBON 6.0 .4508681251 2.5876345040 .4070529028

Ge_8C_{12}

GERMANIUM 32.0 1.5139018204 -1.6017021240 -1.7067478937

GERMANIUM 32.0 -1.8406878963 -1.5577146367 -1.9546366902

CARBON 6.0 -.0659726021 -.6654518149 -2.4181890710

CARBON 6.0 2.1477938074 -.0020432112 .6647011537

CARBON 6.0 .4498202192 2.6204305616 .4098284670

Sn_8C_{12}

TIN 50.0 1.9030985145 -1.8613422649 -2.0251536432

TIN 50.0 -1.7891100376 -1.8048543891 -1.9030349949

CARBON 6.0 .0378767526 -.6383256870 -2.6043459653

CARBON 6.0 2.3341359423 .0023409940 .6380446805

CARBON 6.0 .4594351979 2.6055536774 .4099426848

d D_{2h}

Si_8C_{12}

SILICON 14.0 .0000000000 2.5069667330 -2.4758683579

SILICON 14.0 -2.0159764165 .0000000000 2.4562581285

CARBON 6.0 -2.2271124123 .0000000000 .6028620530

CARBON 6.0 .0000000000 -.6414837498 -2.4584753441

CARBON 6.0 .0000000000 -2.6464440935 -.6038382152

Ge_8C_{12}

GERMANIUM 32.0 .0000000000 2.5914352541 -2.5542715779

GERMANIUM 32.0 -2.1581351664 .0000000000 2.5371502741

CARBON 6.0 -2.3407535032 .0000000000 .6035794387

CARBON 6.0 .0000000000 -.6358001998 -2.5428450726

CARBON 6.0 .0000000000 -2.7089715689 -.6040527553

Sn_8C_{12}

TIN 50.0 .0000000000 2.7949635620 -2.7345325228

TIN 50.0 -2.3649643399 .0000000000 2.7231187797

CARBON 6.0 -2.5727569684 .0000000000 .6054089755

CARBON 6.0 .0000000000 -.6318494280 -2.6569537579

CARBON 6.0 .0000000000 -2.9436686666 -.6054472592

MCSCF (8,8) optimized geometries:

a T_h

Si_8C_{12}

SILICON 14.0 1.5376486991 -1.5393914780 -1.5439440827

CARBON 6.0 .0000000000 -.6752683316 -2.2670324329

Ge_8C_{12}

GERMANIUM 32.0 1.5906092780 -1.5966925975 -1.6000840726

CARBON 6.0 .0000000000 -.6682670768 -2.3397293139

Sn_8C_{12}

TIN 50.0 1.822586 1.822586 1.822586

CARBON 6.0 0.00000 0.641930 2.587048

b T_d

Si_8C_{12}

Ge_8C_{12}

Sn_8C_{12}

c D_2

Si_8C_{12}

SILICON 14.0 1.4222152004 -1.5682583592 -1.6513226343

SILICON 14.0 -1.8089774725 -1.4464750635 -1.9316973553

CARBON 6.0 -.0746556150 -.6725372815 -2.3506933716

CARBON 6.0 2.0580784609 .0200578012 .6697290571

CARBON 6.0 .4441102573 2.5901731744 .4140916035

Ge_8C_{12}

GERMANIUM 32.0 1.4931082254 -1.5936564210 -1.7571177022

GERMANIUM 32.0 -1.8363779118 -1.5755934479 -1.9413459302

CARBON 6.0 -.0831557577 -.6568920212 -2.4822997513

CARBON 6.0 2.1032394385 -.0188802377 .6618426979

CARBON 6.0 .4346720428 2.5965074106 .4261300488

Sn_8C_{12}

TIN 50.0 1.8928276262 -1.8778874910 -2.0280701618

TIN 50.0 -1.7691394320 -1.8172583608 -1.8958593449

CARBON 6.0 .0470767321 -.6368216405 -2.5882971629

CARBON 6.0 2.3144696696 .0048341509 .6358241626

CARBON 6.0 .4534645256 2.6067521464 .4164415037

d D_{2h}

Si_8C_{12}

SILICON 14.0 0.0000000000 2.5068567447 -2.4758375486

SILICON 14.0 -2.0162703793 0.0000000000 2.4563043903

CARBON 6.0 -2.2269435072 0.0000000000 0.6033705940

CARBON 6.0 0.0000000000 -0.6414294394 -2.4584868108

CARBON 6.0 0.0000000000 -2.6464873054 -0.6038315139

Ge_8C_{12}

GERMANIUM 32.0 2.5541997031 -2.5922701325 0.0000000000

GERMANIUM 32.0 -2.5378326988 0.0000000000 -2.1909107786

CARBON 6.0 -.6035497974 0.000000000 -2.3654757549

CARBON 6.0 2.5383535464 .6305325057 0.000000000

CARBON 6.0 .6039457773 2.7091963307 0.000000000

Sn_8C_{12}

TIN 50.0 .0000000000 2.7998312777 -2.7340081524

TIN 50.0 -2.4053228699 .0000000000 2.7239144412

CARBON 6.0 -2.6051491135 .0000000000 .6053502348

CARBON 6.0 .0000000000 -.6263984078 -2.6445859968

CARBON 6.0 .0000000000 -2.9509433784 -.6053463166

CHAPTER 6: CONCLUSIONS

In this chapter the overall results of each project presented in this dissertation are summarized including: the theoretical determination of the minimum energy reaction pathway for the titanium catalyzed hydrosilation reaction (Chapters 2 and 3), the theoretical determination of the minimum energy structures and lowest energy state of a set of main group analogs to metallocarbohedrenes (Chapter 5), a new method and implementation for computing integrals involving Effective Core Potentials (Chapter 5), and finally a new graphical user interface for the GAMESS program (Chapter 4).

Chapters 2 and 3 describe the minimum energy reaction path for several titanium catalyzed hydrosilation reactions. In chapter 2, the simplest model system was considered using hydrogen atoms for all substituent atoms. In chapter 3 the experimental systems were considered by adding chlorines and cyclopentadienyl rings to the model system. Overall, the model system presented in chapter 2 was very successful in describing the minimum energy reaction path for the experimental system at a much lower computational cost than the experimental systems.

The most important points along the minimum energy path were described quite well by the model system. In particular, all reaction systems studied exhibit no overall barrier in contrast to the large (>50 kcal/mol) barrier for the uncatalyzed system. This is largely due to the ability of titanium to insert into the carbon-carbon double bond to form a very low energy complex. However, due to titanium's ability to insert into bonds, particularly Si-H bonds, there were several significant differences between the model system and the experimental systems. These differences were mostly a result of the fact that hydrogen does not fully model either the electronic or the steric effects of either chlorine or a cyclopentadienyl ring.

However, with the inclusion of the full experimental substituents, the final reaction system studied is expected to accurately predict the minima and energetics of the experimental reaction system. While the basic reaction path has been determined there are several areas available for future study. These include examining the basic reaction path using dynamics calculations to better determine the barrier heights, particularly in the portion of the reaction path around the first transition state. In addition, other experimental reaction systems could be studied by using substituents other than hydrogen and chlorine and catalysts other than divalent titanium.

While working on the hydrosilation reaction, it became clear that the existing visualization tools were inadequate for visualizing the results of a complex reaction. Chapter 4 describes the MacMolPlt program that was written to fill this need. MacMolPlt provides all users with the ability to visualize complex systems either through static images or via animations of reactions and molecular vibrations. In addition, all of the most useful surface types, including orbitals, total electron densities, molecular electrostatic potentials and density differences, can be computed directly by MacMolPlt. Thus, MacMolPlt greatly simplifies visualizing the output from quantum chemistry applications such as GAMESS, by removing steps such as conversion of the output of one program into the input for the next program. In addition, high quality 2D and 3D graphics are produced through use of graphics libraries built into the system software. Thus, MacMolPlt is freely distributed worldwide without requiring the user to purchase any third party software. This has made MacMolPlt popular for use in teaching chemistry in student computer labs where the budget for software is often quite small. In addition the powerful animation and surface visualization features have proven useful to even the most experienced quantum chemist.

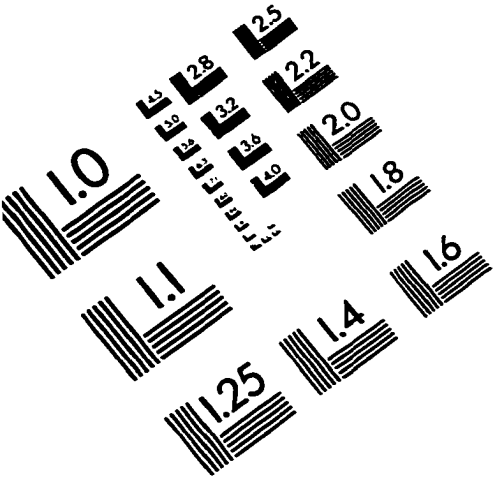
While the existing program performs well, there is plenty of room for improvement. Future enhancements may include an improved graphics engine to improve many areas of visualization. Additional file parsing code to allow output from other quantum chemistry programs such as, Gaussian 9x or ACES II to be read in as easily as it currently is to read GAMESS output. Enhancements to the GAMESS input generator would also be useful to allow users to create complete input files for even complicated calculations. In addition, the code could also be ported to other operating systems to allow more users to take full advantage of MacMolPlt's feature set.

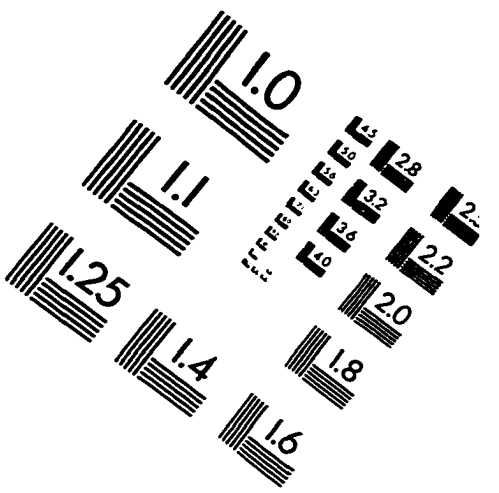
One point illustrated nicely in Chapter 3 was the effectiveness of Effective Core Potentials (ECP's) at reducing the cost of a calculation without reducing the accuracy. However, it also illustrated where improvements were needed. The derivation and implementation of these improvements is described in chapter 5. At the core of these improvements is an improved method for computing the ECP integrals which is not only faster (by a factor of 6), but also scales to higher angular momentum. This is particularly important since it is often necessary to include basis functions with higher angular momenta than the highest occupied atomic orbitals to allow for polarization of bonds and lone pairs. Thus, the inclusion of the f and g-type basis function capability into GAMESS allows more accurate calculations on elements in the fourth and lower rows of the periodic table, particularly for the rare earth elements which include occupied f-orbitals. Inclusion of higher angular momenta is also important in the computation of analytic derivatives of the energy since the angular momentum is increased by one for first derivatives and two for second derivatives. This improvement in the computation of the integrals has also allowed for the inclusion of analytical second derivatives making the computation of vibrational frequencies

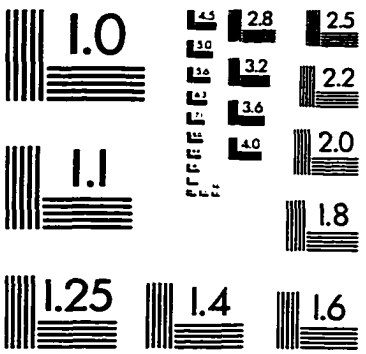
more accurate and much faster. Thus, the inclusion of the improved ECP integral package has greatly improved the performance and usefulness of computations involving ECP basis sets.

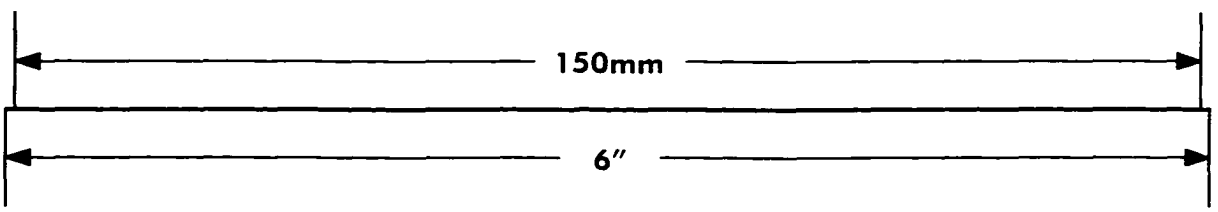
The value of ECPs was also demonstrated in chapter 5 through their use in determination of the various minima for Si, Ge, and Sn analogs of the metallocarbohedrenes (metcars). These systems can be represented by E_8C_{12} (where E = Si, Ge, or Sn) and were generally thought to have a cage-like structure similar to C_{20}, the smallest cage-like fullerene. ECP's allowed the exploration of many possible structures and made computational cost of replacing Si with Ge, or Sn negligible since all three compounds have the same number of valence electrons and the same size basis set. The results of these calculations produced some unexpected minima that do not have a very similar appearance to the expected cage-like structure. In particular, two low energy minima possessing a symmetry of D_2 and D_{2h} were found. In addition one of the expected structures, possessing T_d symmetry, was found to be a high order saddle point. For the tin and germanium compounds, the D_{2h} structure was found to be the global minima. Geometries, energy differences and vibrational frequencies were also reported. Future work will look for these new types of structures in the transition metal based metcars such as Ti_8C_{12}, as well as exploring the stability of each of the structures presented in this study.

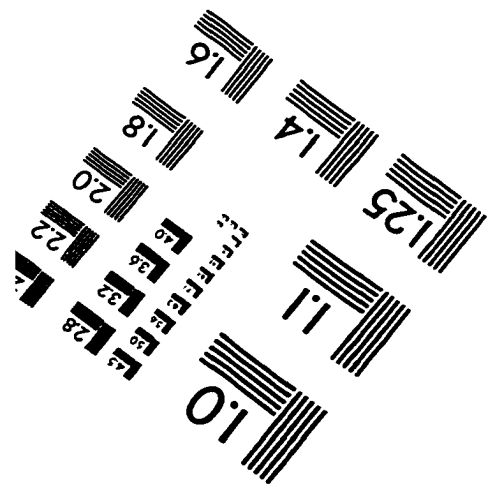
IMAGE EVALUATION
TEST TARGET (QA-3)

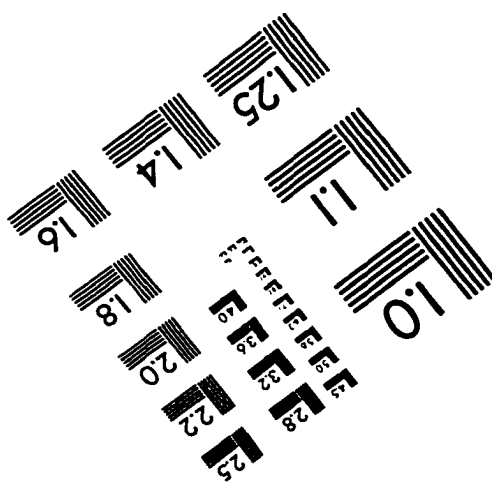












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989